9

A View to the Core

Technological Units and Debitage Analysis

ANASTASIA STEFFEN, ELIZABETH J. SKINNER,
AND PETER W. AINSWORTH

In this chapter we examine available techniques for the analysis of debitage, the waste products generated during the manufacture of flaked stone artifacts. It is our observation that much of the difficulty in understanding debitage literature derives from the form and structure of the units of analysis. Our goal is to distill and make explicit the underlying organization of debitage analyses. To this end, we focus on the units of observation and inference employed in some of the most widely known and popularly used debitage analysis techniques in North America.

Two themes guide the discussion that follows. The first theme is that lithic debitage analysis necessarily falls under the study of reductive technology. The chapter, then, begins with general discussions of flaked stone as a reductive technology and of units in lithic technology. This examination of the relationship between core reduction and basic descriptive units is designed to provide a foundation for the subsequent evaluation of debitage analyses. The second theme is that the validity of analytical techniques is significantly influenced not only by the content but also by the structure of the units employed. The units created to record observations about debitage have important consequences not only for the character of the data generated but also for research outcomes and, sometimes, for the formulation of future research questions. To identify how this interplay is expressed, we examine the intrinsic unit structure of some of the available debitage analysis techniques and evaluate the validity of the techniques in light of unit structure. We have not attempted a review of all techniques, nor a comprehensive guide to the debitage analysis literature (for a recent review, see Shott 1994). Rather, our aim is to provide guideposts so that individual researchers can select techniques that are adequate, appropriate, and feasible for their research. In essence, we examine the method behind the techniques, a necessary first step in evaluation.

FLAKED STONE TECHNOLOGY AS REDUCTIVE TECHNOLOGY

We define technology as "the methods and materials used to manufacture a product." The concept of technology can be broken down into three subdivisions: reductive, additive, and altered. In a reductive technology, products are manufactured solely through the removal or subtraction of material. In contrast, additive and altered technologies include the manufacture of materials through the appending, joining, combining, or compositional alteration of like and unlike materials. Ceramic technology is an obvious example of an additive/altered technology, and lithic technology is an obvious example of a reductive technology (except for lithic heat treatment, which is an altered technology).

In a reductive technology only a single

kind of material is manipulated at a time—in fact, only a single item or mass is manipulated. Furthermore, the only manipulation is the removal of material from that mass (i.e., the core). Finally, the manipulation of that mass is simple in mechanical terms: fracture is the single mechanical process involved. Examples of fracture include flaking, pecking, grinding, abrading, and polishing, with each distinguished by the decrease, respectively, in the size of the fracture. This discussion focuses on the manufacture of stone artifacts through the control of relatively large fractures (i.e., "flaking"). Therefore, we use the term "lithic technology" to refer specifically to flaked stone technology.

Identifying a technology as reductive has several implications for the archaeological units we create to investigate that technology. Given that the archaeological record is dominated by bits and fragments, the units created for objects are frequently constructed to investigate the whole by examining the pieces.[1] In a reductive technology, what does the artifact as a basic archaeological unit represent about some larger whole? What information on technology is contained in the artifact, and how does this compare with artifacts of an additive or altered technology?

Consider the differences between a ceramic sherd and a lithic flake. A sherd is a broken piece of a final product (e.g., a vessel), and as an object, the sherd is created by breakage, or failure, after the completion of manufacture. While a sherd contains information on the materials and manipulations used in manufacture, it represents the final product. In contrast, a flake is rarely a portion of some other final product. As an object, a flake is created as part of the manufacture, not the fragmentation, of a product. The flake itself may serve as a final product, but does not necessarily contain any information about any other product. Thus, the larger whole that a flake represents is not a product, but the manufacture process itself. A flake contains a unique record of the process of manufacture employed in its creation. This remarkably detailed record of the lithic manufacture process is abundant and ubiquitous, and there is no similar artifactual record in ceramic assemblages (except in very rare cases).

It is the simplicity of manufacture that gives debitage assemblages such information potential: reductive technologies leave a record of the manufacturing process on the pieces removed. A flake, unless substantially modified after detachment, retains information on the portion of the core from which it was removed, at the point when it was removed. This information includes not only the morphology of the core prior to removal but also the manipulation that resulted in fracture and detachment. In contrast, the core contains a record of its own final morphology and a record of the final series of removals from the it. The core and removals together contain a complete record of the reduction events.

The simplicity of reductive manufacture has an even more basic implication: reduction determines almost all empirical aspects of these artifacts. Even if pertinent research goals are not explicitly technological, it is hardly possible to describe a piece of debitage without some reference to results of reduction. Indeed, we rely on the redundant characteristics of conchoidal fracture to identify a flake as an artifact rather than merely as a rock. Most dimensions of morphological variation in a piece of debitage describe the consequences of the reductive method of manufacture. These dimensions include the host of "flake characteristics," such as interior versus exterior surfaces, relative thickness or width, platform, terminus, waves of force, flake scars, cortex, etc., as well as the more fundamental measurements of size, shape, and weight. It follows from the identification of flaked stone technology as reductive technology that technological considerations be integral to virtually all of the units we use to describe lithic artifacts. Similar technological observations should be made on all items in the assemblage if a goal is to generate technological data. Consistency in analysis should also extend throughout the unit structures discussed in this chapter. Thus, debitage units should accord with lithic technology units, which, in turn, should be consistent with units of a reductive technology. Viewing flaked stone technology in terms of reductive technology provides an important perspective for understanding and improving the internal coherence among the units we use to describe lithic artifacts.

UNITS IN LITHIC TECHNOLOGY

While particular research questions pursued vary widely, all lithic analyses investigate variation in reduction. That is, unless only raw material characteristics are described, all technological inferences pursued about lithic artifacts will be built on descriptions of reductive variation. Thus, the broad goal of lithic technology research is to make inferences about the kinds of manufacture that produce lithic assemblages and to investigate and/or explain the variation observed in and among lithic assemblages.

Basic observational units designed to measure this variation should be created in terms of reduction. First, they should adhere to guidelines for unambiguous archaeological unit construction. That is, they need to be clearly defined, mutually exclusive, and employ only criteria that can be directly observed. Second, these descriptive units should be explicit technological units. Since nearly all of the relevant variation is a result of reduction, the adequacy of lithic technology descriptions will depend on how well the observational units accord with basic tenets of reductive technology.

Basic units are best defined with simple terminology (i.e., encompassing the fullest range of variation while incorporating the fewest number of assumptions). The key *technological* distinction in lithic artifacts is between cores and removals. We define a *core* as a nucleus from which material has been removed. *Removals* are defined as pieces that have been removed from a core and that themselves show no evidence of further postdetachment reduction.[2] These two categories are mutually exclusive and encompass the entirety of material resulting from reduction. At the minimum, there are two pieces that compose a lithic manufacture assemblage: a core and a removal. And at a maximum, regardless of how large, complex, and variable a reductive manufacture assemblage may be, these two units describe the entirety of lithic materials represented with the same amount of sufficiency as when the assemblage contains only the two pieces, core and removal. Distinguishing further technological variation will proceed from specific research goals.

It is our perception that lithic technology studies do not usually investigate the full range of reduction variation. Many North American debitage analyses are driven by implicit emphases. In particular, there is a research focus on biface production that has influenced not only the direction of research but also the basic units used. This narrowing of the range of core reduction variation to be investigated may be a reasonable and appropriate research strategy, but should be a methodological decision that is explicit in the design of research.

This "biface bias" has developed as a result of several factors. Paleo-Indian studies have been especially influential because the Paleo-Indian archaeological record, with its remarkable, finely executed bifacial forms, captured the fascination of North American archaeologists early on (e.g., Roberts 1935b, 1937b). Later, these bifaces served as a powerful stimulus for the growth of flintknapping both as avocation and as research strategy (see Johnson 1978; also Flenniken 1984). For chronological goals, Clovis and Folsom bifaces have served as exceptional temporal markers for Paleo-Indian (see LeTourneau, this volume), and have thus become paragons in the typological dominance of bifacial forms over nonbifacial lithic materials employed in the construction of culture historical sequences throughout North American prehistory.

Lithic research in the 1970s and 1980s focused on processual questions that led to the creation of production trajectory and reduction sequence models (e.g., Bradley 1975; Callahan 1979; Collins 1975; Flenniken and Ozbun 1988; Muto 1971; Young and Bonnichsen 1985). These models were fre-

		FUNCTION / USE	
		Product (tool)	**By-product** (non-tool)
TECHNOLOGY	**Core**	*biface projectile point*	*blade core*
	Removal	*utilized flake*	*waste flake*

Figure 9.1. Intersection of technological units and functional units; entries in each cell are examples.

quently, although not exclusively (e.g., Bordes and Crabtree 1969; Crabtree 1968; Sheets 1975), focused on biface production. Bifacial reduction is heavily represented in the lithic experimentation that developed subsequent to this production modeling (Magne 1985:25). Finally, biface production is central to much of the research concerned with the organization of technology (see review article by Nelson 1991). This research, inspired in large part by the concepts of expedient and curated technologies proposed by Binford (1977, 1979b), but now substantially augmented and transformed, is currently one of the most popular for lithic studies as applied to well-defined research goals (e.g., Bamforth 1991b; Bleed 1986; Jeske 1992; Kelly 1988; Parry and Kelly 1987; and most recently, an entire edited volume [Carr 1994]).

The problem we see with the research emphasis on bifaces is that it has contributed to a conflation of technology and function in the construction and use of basic lithic units. A ubiquitous example can be seen in the conventional usage of the terms "core" and "tool," especially when used in the context of distinguishing between "core reduction" and "tool production." These seemingly neutral phrases are rich in information content, as a simple breakdown of the phrases demonstrates. First, reduction is contrasted to production—a distinction that makes little sense in a reductive technology. Second, core is distinguished from tool—a contrast that is not technological, but rather pertains to artifact use.

Our concern with these phrases is more than semantic: the basic terminology employed in archaeological lithic studies specifies our most fundamental units, and ultimately can influence our approaches to research. What on the surface is little more than a naming convention actually belies a set of units constructed without accord to reductive technology. When the phrases "core reduction" and "tool production" are used in contrast to each other, they effectively serve to specify the range of reduction variation that will be considered in a study and to what detail the variation will be examined. That is, tool production is the focus—where "tool" actually means biface. All other variation in reduction is subsumed under the "other" category called core reduction. That is, while a category called core reduction should encompass all variation in lithic manufacture, in this dichotomy, tool production is highlighted and all other kinds of reduction are treated as generic.

Deciding whether to call a biface a core or a tool should be a methodological problem solved by selecting or creating units relevant to the specified research goals.[3] As defined above, the central technological distinction is between cores and removals. In contrast, research goals that are functional require units that are defined in terms of use, not manufacture, because the basic question is whether an item was used as a tool. Therefore, the key *functional* distinction is between product and by-product. *By-products* are items that have no evidence of postmanufacture manipulation (use); *products* are items that do. These categories are mutually exclusive, defined wholly in terms of use, and include only properties that are empirically observable. Intent, manufacturing goals, and object form are not necessary criteria.

Figure 9.1 shows these units (core and removal, product and by-product) as dimensions of variability. Clearly, core and tool are units derived from different research domains. Cores can be by-products or products; by-products can be cores or removals. To assume otherwise is to conflate technological with functional research goals. Further, this blurring of functional and technological units betrays a tendency to conceive of artifact description in terms of the products of manufacture rather than the process of manufacture—a perspective appropriate for an additive technology but not for a reductive technology. The biface bias is the archetype of this prejudice.

UNITS IN DEBITAGE ANALYSIS

In this section we incorporate the perspective on lithic technology as reductive technology to describe the unit structure of some of the most popular and widely known debitage analysis techniques used in North America. Deciding which technique or techniques to use requires evaluating the validity and reliability of a technique as well as considering pragmatic factors such as time and labor efficiency. How well a technique performs within each of these areas of consideration depends in large part on how the observational (analytic) units are structured. Thus, we provide the following descriptions of unit structure, with a focus on scale, as important background for understanding the strengths and weaknesses of the techniques as research tools.

We have grouped debitage analysis techniques into three categories according to the scale of observation employed. From least to most inclusive, these three categories are attribute, item, and aggregate. Scale has two important properties: inclusiveness and resolution (see chapter 1). Inclusiveness pertains to the scope or range of phenomena for which units are constructed (Dunnell 1971:145–47), and resolution pertains to the specificity obtained. The relationship between these properties is inverse; as inclusiveness increases, resolution decreases. Inclusiveness determines the breadth of information possible at a given scale, while resolution determines the amount or detail of that information (Figure 9.2). Thus, simply identifying the scale of the units used in a given debitage analysis technique will provide a good deal of information on the form of the data that will be produced.

The use of scale-defined categories is familiar in debitage analysis literature. Our categories are similar, for example, to Ahler's distinction between individual flake and flake aggregate analyses (e.g., 1989b:86–87) and Shott's distinction (1994) between formal and mass analyses (the latter term coined originally in Ahler 1972 [as cited in Ahler 1989b:95]). Here, however, we specify three rather than two scales; our two lower scales, attribute and item, subdivide their single lower scale.

The three categories—attribute, item, and aggregate—serve to organize the descriptions that follow. We explore the role of scale in determining the kind of information gathered using these observational units. We also make note of the implications of scale for several practical considerations (Table 9.1). These are identified as difficulty, expediency, reliability, flexibility, and "generalizability" (Amick et al. 1989:4).

ATTRIBUTE ANALYSES

In attribute analysis, the analytic units are the individual attributes (properties and variables) of removals recorded for each specimen. These techniques record observations at the lowest scale of inclusiveness. The attributes selected can number 1 to n, and usually include some combination of variables and properties of raw material, flake mor-

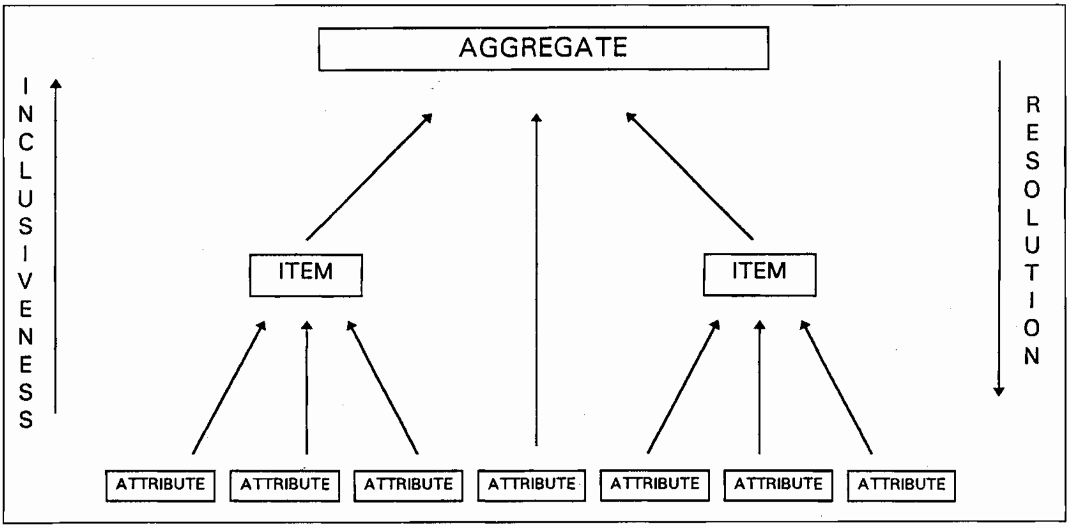

Figure 9.2. Scales of observation used in debitage analyses.

phology, and size. A range of continuous and categorical measurements (i.e., nominal, ordinal, interval, and ratio) may be employed.

Attribute scale analyses were instrumental in the early development of systematic debitage analysis in the 1970s. These early efforts (e.g., Knudson 1973; Phagan 1976; Wilmsen 1970) were exploratory and creative, producing lengthy attribute inventories. The goal was exhaustively to identify attributes that could potentially measure relationships between the control of fracture and patterning in characteristics observable on flakes. The subsequent development of attribute analysis has tended toward refining rather than augmenting early attribute lists, with a pronounced trend toward reducing the number of attributes recorded.

As with all techniques we consider, the inclusiveness and resolution of attribute analysis play a significant role in determining the kind of information created. Because the units in attribute analyses are defined at the least inclusive scale, the resolution of information is extremely high; the observations recorded provide the most detailed information of all the analysis techniques. Usually, the data are generated in a form that maintains a link between specimen and observations, which allows any combination or recombination of attributes to be examined by specimen. Thus, attribute analyses are capable of recording large amounts of information (depending on the number of attributes included) and are highly flexible because, through the selection of individual attributes, they can be precisely tailored to research goals.

This flexibility has certain practical implications as well (see Table 9.1). The independence of attributes allows attribute analyses to be highly generalizable; once generated, data can be used for purposes beyond those defined for the original analysis. These strengths of attribute analysis are also its greatest weakness, especially in practical terms: recording high resolution information results in a technique that is time-consuming and labor-intensive. The difficulty of application (i.e., the degree of familiarity with debitage morphology required to make accurate observations) varies with the actual attributes selected. Once again the flexibility of the technique serves the analyst: it is possible for inexperienced analysts to include only those attributes with which they are familiar.

ITEM (FLAKE) TYPOLOGIES

Techniques at the item level of observation record information in a single observational

TABLE 9.1

Practical Considerations of Debitage Analyses at Differing Scales

INCLUSIVENESS →

← RESOLUTION

	Attribute Analysis	Flake Typologies			Mass Analysis
	Multiple Attributes[a]	Technological Types	Flake Completeness	Cortical Categories	Size Grading
DIFFICULTY: making observations —Expertise required	Variable	High	Low	Low	Low
EXPEDIENCY: creating data —Rapidity of observation	Low	Medium	Medium	Medium	High
RELIABILITY: assessing data —Replicability of observations	Variable	Low	Medium	Medium	High
FLEXIBILITY: manipulating data —Resolution/Amount of information	High	Medium	Medium	Low	Low
GENERALIZABILITY: obtaining data —Broadness of utility	High	Medium	Medium	Medium	Medium

[a] Number and kind of attributes significantly affect difficulty, expediency, and reliability.

step. The analytic units are item categories, usually "characteristic flakes," or individual specimens with specified groups of distinctive attributes. In contrast to attribute analyses, item (flake) typologies record attribute clusters not by attribute but as they co-occur by item.

We recognize three main kinds of item (flake) typologies: cortical categories, technological types, and flake-completeness categories. *Cortical categories* are the simplest: only a single attribute, dorsal cortex coverage, is observed or measured for sorting items into ordinal categories. The historical popularity of cortex categories is likely due to the ease of application (given that only one observation is required) as well as to historical precedence (it is one of the oldest of all debitage units, originally introduced by White in 1963, and subsequently subject to numerous reinterpretations [see discussion by Sullivan and Rozen 1985:755–57]). Because these are essentially single-attribute categories, we do not discuss cortical categories in any further detail.

Technological types are much more complex than cortical categories: they combine multiple observations and require considerable familiarity with debitage morphology. They are most often used by analysts with direct training in the technique. Flenniken's types (1981; Flenniken and Ozbun 1988) focus on variation in flake morphology that is indicative of removal during various portions of modeled reduction trajectories. This focus on sequential variation is readily apparent in the names used for the categories: primary and secondary decortication flakes, early bifacial thinning, late bifacial thinning, edge preparation flakes, and so forth. Another example of a technological typology is provided by Ahler (1986, 1989a; also described in detail by Root [1992]). These types focus on different kinds of reduction, again evident in the type names: bipolar flakes, bifacial thinning flakes, hard hammer flakes, freehand flakes, and so forth.

Finally, the third kind of flake typology is *flake-completeness* as introduced by Sullivan and Rozen (1985). Their categories—complete flakes, broken flakes, flake fragments, and debris—are based not on technologically defined attributes but rather on attributes designed to measure flake breakage. Sullivan and Rozen explicitly designed this new kind of flake typology as an alternative to cortical categories and technological flake types. As is discussed later, the introduction of this typology, as well as the explicitness with which it addressed unit concerns, touched off debates and stimulated research that explored many of the implications of unit structure in debitage analysis addressed here.

To summarize the item-scale units, flake types are constructed at the middle level of inclusiveness and resolution. Compared to attribute analyses, detail is diminished because information is recorded as grouped rather than as individual attributes. Typically, the amount of information is lessened as well because fewer observations are made per item; cortical categories are at the extreme low end with only one attribute recorded, the flake-completeness types are intermediate, and the technological types are at the high end. When flake typologies are used as a sorting procedure, and information is not recorded by individual (i.e., numbered) specimen, it becomes impossible to manipulate information except by frequency of flake type.

However, the use of item-scale and combined-attribute observations affords significant practical advantages. Most important, it is possible to record a relatively large amount of technological information in a manner that is, potentially, much more rapid than for attribute analysis. Using flake types can be so efficient that they are well suited to in-field analyses. However, determining whether this gain in efficiency is worth the loss in resolution of information requires a consideration of validity, not simply practicality.

AGGREGATE (MASS) ANALYSES

Techniques at the highest scale of inclusiveness are those that record observations on entire aggregates of debitage. Information is recorded for groups of debitage rather than for individual items. Thus, unlike techniques at the two lower scales, mass analyses record attributes of an aggregate rather than attributes of an item (flake). Aggregate analyses vary in complexity. Ahler's work (e.g., 1986, 1989a, 1989b) represents both the most complex and the most comprehensive application of mass analysis, while Patterson's work (e.g., 1990) exemplifies a less complicated formulation. Despite differences, the sorting of debitage aggregates by size-grade is common to all aggregate analyses. The analytic units are frequency counts of items by size-grade and summary weights of the size-graded subaggregates. Given the high level of inclusiveness, the resolution is necessarily low: the information obtained includes only a small number of attributes, such as count or weight by subaggregate.

The great advantage of an analytic unit defined at the highest level of inclusiveness is that it enables extremely rapid artifact processing. Where only size is recorded, entire assemblages can be processed in a single step. Further refining of the information obtained—for example, by separating debitage by raw material—adds to the time required. Because the main observations are made by using an external instrument (e.g., graded sieves, balance) rather than human perception, the technique requires almost no expertise, employs unit definitions that are explicitly defined and standardized, and therefore is presumed to be highly reliable. Clearly, however, the amount of information recorded is reduced to only a few attributes. In contrast to attribute analyses and flake typologies, there is no direct observation of flake morphology.

To summarize the three scales of debitage analysis, attribute analyses record the most information and are the most flexible because information is contained in packages that can be recombined and manipulated most readily. However, recording all this information is costly in terms of time and labor. Item (flake) types occupy an intermediate position of information content, time required, and flexibility of data manipulation. Aggregate (mass) analyses are the most rapid for processing debitage, but the amount of information obtained is quite small.

Describing the scale of observations used in the various techniques makes clear the determining role that inclusiveness and resolution play in the form of information recorded. In practical terms, these aspects of unit structure are responsible for the inverse relationship between speed of data recording and the amount of information recorded. The balance between speed and information is, of course, what is described by the term "efficiency." However, it is important to distinguish between efficiency in practical terms and the adequacy of analysis for specified research goals. For determining the validity of a technique for producing sufficient and relevant data, there is no *a priori* "most efficient" technique. The assessment of validity, or research performance, requires a more complex evaluation.

EVALUATING DEBITAGE ANALYSIS TECHNIQUES

In describing the unit structure of the various techniques, we examined how structure determines the form of information obtained. An evaluation of the utility and adequacy of a technique for specified research applications builds on these descriptions and entails considering how inferences can be constructed using the information gathered. In unit terms, creating inferences involves constructing synthetic units. Just as the structure of analytic units determines the kind of information, the structure of synthetic units constrains the kind of inferences possible.

For the remainder of this chapter, we explore how the relationship between analytic and synthetic units affects the validity of the debitage analysis techniques. For each of the debitage analysis categories, we discuss how analytic unit structure affects validity and examine implications of how synthetic units are constructed from analytic units. We begin with a brief summary of the role of units and the terms of unit evaluation.

ANALYTIC AND SYNTHETIC UNITS

Analytic units are the observational units used to describe empirical properties being measured. Synthetic units organize these analytic observations into inferential aggregates. Thus, "analytic" and "synthetic" describe two different roles that units play in a research design. Several relationships exist between analytic and synthetic units (see Dunnell 1971; also chapter 1). First, analytic units are used for observation, while synthetic units are used in inference. Second, when strictly defined, there is a hierarchical relationship: synthetic units are constructed at a higher scale of inclusiveness than the analytic units from which they are formed. Third, the role of a given unit is relative: the same unit may be analytic in one study while synthetic in another. Problems arise when this swapping of roles occurs in the same analysis.

VALIDITY AND RELIABILITY

Validity and reliability describe two distinct realms of unit evaluation. Validity concerns the relevance of units to the goals of research, while reliability describes the consistency and replicability of measurement (Amick et al. 1989; Carmines and Zeller 1979). The focus of our discussion is entirely on validity.[4] The contrast with reliability, however, is a useful illustration of validity. Reliability describes the redundancy of actual measures or observations, whereas validity describes the degree to which those measures are relevant. Reliability concerns the performance of an instrument in repeated trials, under varying conditions, or with different investigators. Validity addresses the selection of the instrument, and the units that compose it, for specified research goals.

The broad perspective of core reduction, as outlined above, provides the context in which we evaluate the validity of debitage analysis techniques. To restate, the primary purpose of a debitage analysis technique is to describe variation in debitage that was produced by variation in core reduction. Here we consider validity by evaluating how well a debitage analysis technique describes reductive variation. Given the descriptions of unit structure provided earlier, we assess the effect of that analytic structure for the construction and performance of synthetic, or inferential, units.

For example, debitage analyses historically have been dominated by the investigation of three sources of variation in reduction: force application (e.g., hard versus soft hammer, percussion versus pressure), reduction stage or intensity, and technique or kind of core reduction (e.g., bifacial, bipolar, blade, and generalized core reduction). Our discussion of validity pertains to all of these research goals: we examine how the organization of observational units facilitates the empirical description of debitage used for making inferences about these sources of core reduction variation.

Throughout the development of research in lithic technology, experimentation has been the most important means for establishing validity. First, experimentation provides an invaluable opportunity for observing descriptive aspects of debitage that later are codified as analytic units. This productive relationship between experimentation and unit creation is possible in part because of the relative simplicity of reductive technology and the constancy of fracture mechanics.

Experimentation is also used to test the validity of specific techniques.[5] This second use of experimentation has more potential for problems because it involves synthetic units of inference in addition to analytic units of description. The difficulties associated with this complex use of experimentation are well illustrated in debitage attribute analyses, to be taken up in the next section. Thus, we turn now to the evaluation of specific techniques, beginning at the highest resolution.

ATTRIBUTE ANALYSES

The central validity concern for attribute analyses is the selection of attributes to record. Indeed, this is the most fundamental concern for all debitage analyses. In spite of, or perhaps as a result of, the many studies conducted at this scale of observation, there is no consensus on which flake attributes are presumed relevant for describing variation in reduction or on which attributes have been demonstrated to perform well (Ahler 1989b; Shott 1994). As noted above, experimentation has played an important role in the development of attribute analysis, where the primary means for "discovering" and testing for relevancy has been through the use of controlled and replicative experimentation. Controlled experimentation has tended to focus on the relevance of single or few attributes of debitage morphology (e.g., Dibble and Whittaker 1981; Speth 1981), while replicative experimentation has frequently examined numerous attributes, including both morphology and size attributes (e.g., Amick et al. 1988; Baumler and Downum 1989; Henry et al. 1976; Ingbar et al. 1989; Magne 1985; Magne and Pokotylo 1981; Tomka 1989).

Given the potential for generating overwhelmingly large and complex data sets when multiple attributes are used, it makes sense that multivariate statistical analyses were introduced early in the development of attribute analysis (e.g., Burton 1980; Chandler and Ware 1976), and have become increasingly common (e.g., Amick et al. 1988; Odell 1989; Root 1992). In brief, the statistical techniques are used to group relevant attributes or to discriminate relevant from ineffectual or redundant attributes. Although these studies have obtained mixed results for creating effective attribute clusters, they have provided a mass of individual observations that are useful as a basis for the selection of attributes. Thus, experimentation has succeeded in contributing to the most important challenge for attribute-scale analyses: deciding what aspects of debitage variation to describe.

The use of multivariate grouping procedures and experimental replication to validate analysis techniques illustrates a complex unit problem. When experimentally replicated debitage is used, the aggregates of debitage represent some specified "bounding," or subset, of reductive variation that is focused by research interest. For example, the debitage aggregate could represent all of the debitage produced during some defined stage of reduction or the debitage produced from the reduction of a core from beginning to end. For the sake of simplicity, we call these replication aggregates "reduction sets." The problem lies not in using the experimental reduction set as a unit within which variation is examined, but in employing it as a conceptual or statistical "control unit." Multivariate classification procedures, such as discriminant analysis, require that unknown items be compared to a group of known items that was used to create the statistical definitions for group membership; when applied in experimental debitage analysis, the reduction set (or sets) serves as that "known." We are concerned with the validity of units constructed in this manner. The use of reduction sets as control groups effectively specifies that the bounded group designated as the control (e.g., all of the debitage produced during a bifacial reduction stage) is the targeted synthetic unit. That is, if we construct or evaluate our descriptive units by comparison to experimental assemblages, then we are implicitly designating these reduction subsets—and unmixed assemblages—as our desired synthetic (inferential, interpretive) unit.

We emphasize that our concern is not with the sufficiency of the reduction set as a behavioral analog nor with its equivalence to archaeological assemblages subject to formation processes unrelated to manufacture (e.g., Ingbar et al. 1989). Rather, the use of the reduction set as a synthetic unit presumes that our unit design specifies *that aggregate* as the scale of our synthetic unit. The eventual goal of experimental analyses is to make inferences about archaeological deposits, and it is unclear how this unit structure can be valid, given the technological mixing contained in most archaeological assemblages. The mixing is not a problem to be overcome. Rather, technological mixing within assemblages is an intrinsic part of the archaeological record—our units should reflect this.

The issues of how experimental debitage units should be constructed and used and how technological mixing within assemblages should be viewed in the design of research are relevant not only in attribute analyses but also for experimental debitage studies defined at all scales of observation. These concerns are highlighted in attribute analyses because the use of statistical procedures lends an air of objectivity to the process and also because attribute analyses are often used to inform or test debitage analysis techniques at the higher scales of item and aggregate.

FLAKE TYPOLOGIES

Technological Flake Types

Two of the most frequent criticisms of technological typologies are (1) they are subjective and unstandardized, and (2) they suffer from inferential biases (see discussion in Sullivan and Rozen 1985; also Ahler 1989b; Ingbar et al. 1989). Both of these problems stem directly from the unit structure. Thus, considering units identifies the source of the problems and also suggests potential remedies. In the first case, the cause is the form and definition of the analytic units; in the second case, the difficulties arise due to overlap between analytic and synthetic units. We take up each problem in turn.

Flake types are the analytic units used in technological flake typologies. Usually these units were created subjectively from the recognition (i.e., discovery) of patterned variation in debitage observed from replication experimentation (e.g., Flenniken 1981). The primary goal of unit description was to capture the important characteristics for group membership—creating precise definitions was less important because training rather than documentation served as the primary means for sharing type identification. Thus, at their worst, flake types are ambiguously defined. They are polythetic empirical groupings that highlight modal tendencies rather than rigorously specifying the criteria of group membership.

Fortunately, this problem can be easily corrected: all that is required is improvement in how the units are defined. Technological types should be constructed using explicit and standardized definitions that systematically and comprehensively specify the attributes that belong with each unit. The types need to be mutually exclusive with regard to the combination of attributes necessary for membership, yet still be free to vary for other attributes. While experience will always be important for effective identification of attributes, explicit definition will make the data produced from typological analyses more generalizable. Recent work with technological types demonstrates this kind of improvement in the clarity and adequacy of type definition, as exemplified by Root's type definitions and lucid discussion of technological typology (1992:80–89).

The second problem area concerns how the flake types are used in inference. Specifically, analytic units and synthetic units become merged, creating bias in data collection as well as interpretation. Here it is useful to consider some of Sullivan's and Rozen's criticisms (1985) of traditional typologies by restating their arguments in terms of unit structure.

First, Sullivan and Rozen argue that analytic units for describing variation in debitage "should not be linked *a priori* to specific conclusions about lithic technology" (1985: 758). In other words, units in debitage *analysis* should describe variation in core reduction without automatically synthesizing the implications of that description. Second, they propose that inferences about technological origins of debitage should be made at the scale of assemblage rather than individual item. When considered in terms of unit structure, the latter point argues that where analytic units are defined at the lower scale of item, the synthetic units should be constructed at the higher scale of aggregate. In other words, when the relationship between analytic and synthetic units becomes blurred both in definition and in scale, technological typologies are then structured so that the unit of inference is employed as the unit of observation (Shott 1994:77). The technological type "biface thinning flake" serves as an example. Clearly, the name of the unit identifies the kind of core reduction that produced the removal. If the research goal of the analysis includes identifying whether bifacial reduction took place, then typing a specimen also answers the technological question being asked.

The potential for circularity in logic

(Teltser 1991:366) undermines the descriptive capacity of technological types and contributes significantly to their interpretive "subjectivity."[6] Thus, technological types may not be valid descriptive units if applied to debitage derived from kinds of core reduction that are outside of the range considered when the types were originally constructed. For example, the performance of most technological typologies for describing removals generated during bipolar reduction must be questioned if this kind of reduction was not originally considered.

The solution to the conflation of observation and inference in technological typology may not be easily obtained. Much of the inferential power of technological types is achieved as a result of their subjective derivation, which plays a significant role in the "directness" of their link to technological variation (Ahler 1989b:87; Root 1992:88). Despite the existence of methodological problems, technological types do seem to have validity for identifying distinctive patterns of attributes that correspond with the differences in reduction targeted in specific research studies (e.g., Hayden and Hutchings 1989; Magne 1985). Experimental research, especially if broadened to include a wider range of core reduction variation, is one obvious avenue for evaluating the type definitions and refining the selection of attributes included in those definitions.

Flake-Completeness Typology

Sullivan's and Rozen's flake typology (1985) was constructed explicitly to overcome problem areas identified for technological and cortical types. The flake-completeness typology offered certain improvements in unit structure and content: explicit and nonpolythetic analytic unit definitions, a simple and unambiguous specification of criteria for sorting debitage (including only criteria that describe morphological attributes relating to flake breakage), and the specification of the scale of both the analytic and synthetic units. Their typology made an important contribution to the development of debitage analysis: their proposal brought unit concerns to the forefront of debitage analysis and stimulated a wave of evaluation. Indeed, no other debitage analysis technique has received such swift and focused response. Their typology was immediately questioned on theoretical and inferential grounds (Amick and Mauldin, eds. 1989; Ensor and Roemer 1989), and has been subjected to a number of experimental applications (e.g., Baumler and Downum 1989; Bradbury and Carr 1995; Ingbar et al. 1989; Kuijt et al. 1995; Mauldin and Amick 1989; Prentiss and Romanski 1989; Tomka 1989; many of these experimental studies appear together in an edited volume [Amick and Mauldin, eds. 1989] that was compiled largely in response to Sullivan's and Rozen's proposal of the typology).

This outpouring of critical and constructive response to the flake-completeness typology explored or directly challenged the validity of the technique for describing debitage variation relevant to core reduction variation. The intensity of the response was due to the lack of empirically grounded, warranting arguments for validity in Sullivan's and Rozen's original presentation. In other words, Sullivan and Rozen did not satisfactorily explain why their breakage categories should bear any relationship to the reduction variation they targeted (i.e., a comparison of core reduction and tool production). The sparse attribute descriptions included in their category definitions further exacerbated the problem.

Despite these criticisms, Sullivan and Rozen accomplished an important part of their purpose in proposing an alternative to the traditional typologies: they offered a typology that was significantly better organized and rigorous in structure. The unit structure of the flake-completeness typology is explicit and employs an internally coherent set of criteria that are mutually exclusive and encompass the full range of expressed variation within the dimensions employed. The specification of the criteria is explicit, and thus designed to facilitate objectivity and standardization among researchers. Furthermore, their identification of the scalar relationship between the analytic and synthetic units and the specification of synthetic units at the scale of assemblage are thoughtful contributions.[7]

To return to a comparison with the other item (flake) types, it appears that the flake-completeness typology, in solving many of the concerns raised for traditional typologies, abandoned many of the strengths of technological typology. Overall, technological typologies were explicitly derived from and supported by empirical observation of debitage variation based on experimental replication and fracture mechanics research. In contrast, the flake-completeness typology was introduced with little prior external validation. Interestingly, the experimentation generated in response to the typology has provided a base of empirical evaluation that, now, approaches or exceeds other techniques.

An obvious way to maximize the benefits of each kind of typology is to combine the explicitness and strict adherence to observation contained in the structure of flake-completeness typology with the technological content and multidimensional approach of technological typology. Much of the ease, and part of the "objectivity," of Sullivan's and Rozen's typology results from the paucity of information recorded during the sorting process. In this regard, their approach yields little more information than do the cortical categories. The decision whether to employ their typology as it was originally formulated must take into consideration this limited information potential.

AGGREGATE (MASS) ANALYSES

The two most significant concerns about the validity of aggregate-scale analyses have been identified and discussed by Ahler (1989b: 87–93). The first is the "potential lack of clear linkage between the data sets or variables recorded and behavioral variation in the archaeological record" (1989b:88–89). The second pertains to the validity of applying the technique to mixed assemblages. Clearly, the first concern is the issue central to the validity of all debitage analyses. Ahler has emphasized it for mass analysis because the relationship of the attributes measured in mass analysis to the subsequent construction of synthetic units is neither intuitively apparent nor derived from the more customary item- and attribute-scale observations. The explicit discussion that Ahler offers for developing an "explanation for why mass analysis or aggregate analysis *works*" (1989b:89, emphasis in original) contributes substantially to establishing abstract and empirical validity for this technique.

To paraphrase Ahler's points, the general ideas used to support the validity of mass analyses consider how the nature of reductive technology pertains to size and weight as technological units and how variation in reduction (specifically, load application) can be inferred from relatively simple combinations of measurements (of size and shape) undertaken on debitage aggregates. Clearly, these arguments for the validity of the units derive explicitly from a view of lithic technology as reductive technology. Evaluating whether the particulars of these arguments are sufficiently tested or supported by the empirical observations generated can be undertaken by examining the data that he presents (as well as others, e.g., Baumler and Downum 1989; Patterson 1982, 1990; Root 1992; Stahle and Dunn 1982; see also Shott [1994] for a lengthy discussion of specific aggregate analyses).

The second validity concern, the performance of the technique for examining mixed assemblages, is considered by Ahler to be "relatively unique to the aggregate approach" (1989b:89). We do not agree, for we consider the problem of mixed assemblages to be of similar import for all the analyses undertaken at all scales. What is unique is the relationship of aggregate-scale analytic units to aggregate-scale synthetic units: they are constructed at the same scale, and no hierarchical relationship pertains. An important result of constructing units at this high scale of inclusiveness is that it has required researchers to more carefully consider the implications of assemblage as the inferential unit. As discussed above for attribute analyses, use of experimental reduction sets as analogs for archaeological assemblages is problematic when attempting to establish the validity of debitage analysis techniques; these problems are simply more obvious for aggregate-scale analyses.

CONCLUSIONS

Our goal has been to clarify the organization of debitage analysis to enable an evaluation of the techniques for determining appropriateness for technological research goals further specified in individual research programs. Both validity and practicality are relevant concerns for selecting a technique best suited to the application. Our discussion should make it plain why there is no best technique, since these decisions must incorporate a variety of considerations.

If there were no practical concerns, then detailed attribute analyses would offer the greatest potential for information recovery since all other techniques use units that can be formed from attribute-scale descriptions. Probably some combination of the techniques will characterize the most productive and practical approach for each research application.

We initially identified two themes, the relevance of reductive technology to unit construction, and the influence of unit structure in the validity of debitage analysis techniques. We examined the latter theme by describing the unit structure and then by demonstrating the role of that structure in key issues of validity for each set of techniques. Two issues emerge as central to validity for all of the analysis techniques we have discussed. These issues are (1) the selection of specific analytic units for describing relevant variation in debitage and (2) the construction of synthetic units to accord with the inferences sought for archaeological debitage assemblages. The former concerns observations at the finest resolution, the latter concerns inferences at the largest scale of application.

Improving the techniques will require continued attention to the two validity issues stated above. In pursuing this goal, we are well served continually to reconsider debitage analysis in terms of reductive technology. Thinking of stone tool production as core reduction suggests novel possibilities for reformulating our analytic and synthetic units. For example, the problems of using assemblage as a synthetic unit when applied to archaeological debitage assemblages that are technologically mixed may be reduced if we consider the core rather than the assemblage as a synthetic unit. Instead of treating assemblages as samples of the debitage aggregates produced by complete reduction trajectories or as imperfect proxies of lithic final products (tools), we might consider individual pieces of debitage as representatives of the core at the point in reduction when the removal occurred. This approach requires only a subtle shift in attention, but may have substantial consequences for how we construct units for debitage analysis. From this perspective, debitage represents the manufacture process by representing the state of the core. Our inferences about variation in reduction would focus more directly on variation in how cores were reduced. Likewise, focusing more directly on how cores are reduced rephrases reductive variation as core variation, and the range of variation in core reduction examined in debitage analysis would automatically broaden.

This could likewise improve our use of experimentation. For example, with this view to the core, the unit of comparison in experimental replication studies shifts from idealized debitage assemblages to individual removals, each one representing differing core states during a known process, or trajectory, of reduction. The mixing of archaeological debitage assemblages would pose less of an obstacle to inference.

These suggestions serve two purposes. The first is to demonstrate that the available debitage analysis techniques do not represent the totality of possible descriptive techniques and analytic approaches. The units used in lithic analysis need not be viewed as static. The second is to illustrate that reconsidering debitage analysis from the most inclusive perspective of reductive technology not only serves as a check on the internal coherence of basic units but also affords creative potential for constructing new units.

Four major concepts were introduced in this chapter. The first is that the entirety of material resulting from lithic reduction is composed of cores and removals. The second is that a core and its removals together contain a complete record of the reduction event. The third is that the construction of technological units must be research driven. The fourth is that all units in lithic technology describe variation in core reduction. If these four concepts are adopted by lithic analysts, we can begin to develop a universally understood body of basic knowledge that will allow us to move forward.

ACKNOWLEDGMENTS

We would like to thank several people for providing comments and important suggestions on earlier drafts of the chapter: William Dancey, John Fagan, Bruce Huckell, Robert Jackson, Phil LeTourneau, Steven Mack, Kristen Martine, Gerry Raymond, Mark Slaughter, Alan Sullivan, and David Vaughan. In particular, we thank Ann Ramenofsky; her assistance and patience cannot be overstated.

Notes

1. Schiffer discusses the relationship between fragments and wholes with regard to formation processes. In describing his Completeness and Fragmentation Indices (1983:686–89), his comparison of stone artifacts to ceramic and glass artifacts clearly exemplifies how the kind of technology is relevant for archaeological measurement.
2. It is common to further partition removals by whether they bear sufficient measurable attributes deemed relevant in a debitage analysis (i.e., "flakes" do, whereas shatter, angular debris, and chunks do not).
3. Kelly (1988) offers a similar discussion. However, he is concerned with the different "roles" of a biface in the organization of technology, whereas we are discussing categories for biface in the organization of research methodology.
4. Although not considered here, reliability studies are well represented in the lithic analysis literature (e.g., Beck and Jones 1989; Fish 1978; Larralde 1984; Nance 1987; see Shott's discussion [1994:74–75] of measurement error).
5. The use of experimentation for evaluating validity may account for a terminological incongruity: in the debitage literature, the term "reliability" is frequently used to mean what we consider to be validity. For example, it is common in the literature to indicate that a technique is "reliable" for diagnosing the presence of bipolar reduction or that platform lipping is a "reliable" attribute for determining that a soft hammer was used. However, as the terms are defined in this volume, the above are examples of validity, not reliability. Experimentation may be an origin of this blending of reliability and validity, because when a technique is "tested" for accuracy, successful outcomes can be readily conceived of as evidence for the "reliability" of the technique. However, unless the test measures only replicability, the term "reliability" is likely being used in a colloquial sense rather than the more restricted definition employed here. Also, reliability is sometimes used in a similar sense in certain statistical applications, such as multivariate discrimination and clustering routines. This may have contributed to the use of the term in lithic analyses, especially given the association of such statistics and experimental debitage studies.
6. This is, in part, the point that Sullivan and Rozen (1985) were making in their call for "interpretation-free" categories.
7. Nonetheless, certain misunderstandings and disagreements have arisen concerning their treatment of analytic and synthetic units. We believe that much of the debate surrounding inference and methodology contained in the 1989 comments and response articles (Amick and Mauldin 1989; Ensor and Roemer 1989; Rozen and Sullivan 1989; Rozen and Sullivan 1989b) is largely a complex miscommunication among the authors concerning different understandings of the construction of synthetic versus analytic units.