VALIDITY IN LITHIC DEBITAGE ANALYSIS:
AN EXPERIMENTAL ASSESSMENT COMPARING QUARTZITE TO OBSIDIAN

by

DONNA M. MORRISON

B.A. (Honours), Trent University 1990

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS
in the faculty
of
Archaeology

© Donna M. Morrison 1994

SIMON FRASER UNIVERSITY

May 1994

All rights reserved. This work may not be reproduced in whole
or in part, by photocopy or other means, without permission of the author.

APPROVAL

NAME Donna M. Morrison

DEGREE MA

TITLE VALIDITY IN LITHIC DEBITAGE ANALYSIS: AN EXPERIMENTAL ASSESSMENT COMPARING QUARTZITE TO OBSIDIAN

EXAMINING COMMITTEE:

Chair Dr. Jonathan Driver

Dr. Brian Hayden
Senior Supervisor

Dr. Jack Nance
Professor
Simon Fraser University

Dr. David Pokotylo
Associate Professor
Department of Anthropology and Sociology
University of British Columbia

Date Approved: May 17, 1994

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis or dissertation (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this thesis for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Title of Thesis/Dissertation:

VALIDITY IN LITHIC DEBITAGE ANALYSIS: AN
EXPERIMENTAL ASSESSMENT COMPARING
QUARTZITE TO OBSIDIAN

Author:

Signature

Donna M. Morrison

Name

May 17, 1994

Date

ABSTRACT

The purpose of this thesis is to assess the validity of three methods of analysis commonly used in the interpretation of lithic debitage. Obsidian and quartzite are used in replicative experiments to explore the nature of raw material variability and the effects that this factor can have on patterning in flaking debris. It is argued that each type of raw material will produce its own unique debitage profile and that methods designed and tested with one lithic material will therefore be unable to accurately identify the technological origins of the debitage from another material type.

The methodology employed includes research into external sources of information such as petrology and fracture mechanics, as well as internal knowledge gained from archaeological experimentation. Replications of various core and biface reduction activities are carried out so that the waste material from these knapping events can be analysed according to the criteria set forth by the three methods under evaluation: 1."Individual Flake Attribute" analysis based on Magne's flake scar method, 2. Prentiss's modified version of the Sullivan and Rozen flake completeness typology, and 3. Ahler's "Mass Analysis". In order to determine whether these analytical approaches are valid, specific hypotheses are tested regarding their ability to accurately identify the different reduction stages and techniques that produced the experimental flake assemblages.

The results of the analysis support the initial claim that each lithic material will fracture in a unique manner (depending on petrological characteristics and the knapping strategy employed), thereby creating its own individual patterns of debitage. It is shown that these differences between material types are quite pronounced, and that it is not possible to interpret quartzite debitage using methods tested on and developed from obsidian, chert, or basalt. That is, these methods (with the exception of the Mean Flake Weight variable used in Mass Analysis and the Platform Flake/Shatter ratio used in the Individual Flake Attribute analysis) are not valid *externally*, or, beyond the original parameters from which they were designed. Furthermore, when analysing obsidian flaking debris, it is shown that there is often enough variability even in such an "ideal" material to challenge the *internal* validity of these approaches, as well. Consequently, these analytical methods require a greater degree of testing before they can be applied to archaeological samples, since raw material variability is shown to have a greater effect on debitage patterning than previously thought.

It is concluded that it is not necessarily safe to assume that the technological origins of debitage assemblages can be consistently or accurately identified regardless of raw material type. This fundamental theoretical construct, which has been one of the basic assumptions in debitage analysis, needs to be more fully examined before we can produce valid inferences about prehistoric technology. Recommendations are provided regarding ways of overcoming this dilemma, and areas for future research are suggested.

ACKNOWLEDGMENTS

I would like to thank the members of my examining committee for all their help throughout the various stages of my research. My senior supervisor, Dr. Brian Hayden, generously donated his time and energy to carry out the replicative experiments, and provided much assistance in writing and researching this project. Dr. Jack Nance offered a great deal of support and guidance, especially in relation to research design and quantitative analysis of the data. Finally, I thank Dr. David Pokotylo, my external examiner, for his interesting and insightful comments on my thesis project. I greatly appreciate all the contributions that my committee members made.

I am also grateful to the Department of Archaeology for awarding me a Graduate Fellowship, a Research Assistantship, and several Teaching Assistantships. These provided me with much needed financial support and, more importantly, experience in the classroom. Barb Winter also helped me by allowing me to work for her in the Museum during my last semester.

There are a number of other individuals in the Department who also made my years at S.F.U. much more enjoyable, including our two technicians and a wonderful office staff -- many thanks to Andrew, John, Ingrid, Ann, Linda B., and Lynda P.! You made everything go so smoothly. I also appreciate the friendship and support of everyone in the Grad Lab (and beyond): Mike & Darlene, Eldon, Natalie, Hartley, Farid, Rick, Sandi, Kar-el (sorry about your favourite hammerstone!), Tanja, Cathy, and of course Tony (hope I haven't left anyone out).

In addition, there are several people I can blame for getting me started in this field and making me go through all these years of pedagogical purgatory. First of all, Bill Prentiss, who introduced me to the wonderful world of lithics and who was always there to listen and help me develop my research project. I am also indebted to Inga Svenson,

H. McFiggin, and Brandy Jackson for being my role models and mentors -- thank you all for your guidance along the way.

Finally, I would not have made it without the encouragement and forbearance of my friends and family back east (BCTel loves you, too!) and, above all, of my husband. Since I couldn't think of the right words to use, I told him he could write this part himself : "I would like to thank John McIntyre for being perfect in every way (and for writing my thesis for me). Thanks, John, for allowing me to witness daily your brilliance and charm." I couldn't have said it better myself.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE:
INTRODUCTION TO THE RESEARCH PROJECT

This thesis will focus on the issue of whether different forms or types of raw material create different types of debitage. It is unique in its examination of quartzite waste flakes, as compared to the traditionally-employed obsidian, and seeks to understand how such variability can affect the inferences drawn from lithic analysis.

One of the areas of archaeological enquiry that is characterized by rather restricted research parameters is experimental lithic analysis. Over the course of the last two decades many of the abstract interests of this field have been explored, while the practical side has unfortunately failed to expand along with them. In this respect, researchers have spent a great deal of time making higher level, theoretical claims without ensuring that they are backed by proper experimental validation. As a result, many approaches used in lithic analysis (most notably debitage studies) may actually be providing false interpretations not only of prehistoric assemblages, but also of the experimentally replicated data that are used as the basis for further research. The consequences of using flawed conceptual frameworks should be of great concern to all archaeologists; nevertheless, the field is dominated by analytical methods which have yet to be thoroughly validated. As many will admit, such blind acceptance of ideas and methodologies is both dangerous and illogical.

The choice of lithic raw materials used in replicative experimentation provides a clear illustration of how unsound research designs can adversely affect archaeological inference. For the most part, investigations in this area have almost exclusively used either obsidian or chert/flint, despite the fact that these two raw materials are not representative of the majority of stone types employed prehistorically. In fact, Runnels (1985:100) argues that there is an implicit definition of lithic analysis as the study of only flaked chert and obsidian objects. The main reason that these materials are so commonly used in experiments is that they were common in European prehistoric assemblages, and, more importantly, will fracture in a predictable and easily-controlled manner which makes them perfectly suited to experimentation in two respects. Firstly, these types of stone are very easy to learn to knap since they present few problems to the beginner and, secondly, they clearly show all of the important fracture features that identify various aspects of knapping behaviour, such as the presence of a pronounced bulb of percussion indicating the use of a hard hammer. It is no wonder that people have focused on these lithic materials which provide the most easily observable and most straight-forward links to the theoretical concepts being examined and developed.

We run into problems, however, when trying to look for technological patterns in other materials, especially coarse-grained or non-vitreous ones which will produce different types of flakes and fracture patterns as a result of their unique petrographic and mechanical qualities. I encountered this dilemma while searching for a method to use in the analysis of quartzite debitage from a lithic outcrop site in Wyoming, known as the "Big Buzz" site (48BH932). It gradually become apparent that most approaches designed for debitage analysis had been developed either through testing obsidian or chert, or else had never been tested at all. As such, I was not convinced that they would provide accurate or valid characterizations of the assemblage being studied.

Since the most effective way of relating experimental results to the archaeological record is interactively (Amick et al 1989:1), I decided to use the same lithic material found at our site in Wyoming in replicative experiments which would test the validity of three prominent methods of debitage analysis. Even though quartzite was used by humans all over the world since the time of the first stone tool manufacture, general information concerning its use is very scarce in the archaeological literature. I found that it is even more scarce in the research and publications of experimental lithic analysts; in fact, I was only able to come across five studies published in the last 20 years which focused on quartzite in their knapping experiments. Of the earlier studies, Behm and Faulkner (1974) very briefly examined the effects of heat treating on Hixton quartzite, while Toll (1978) looked at use-wear on quartzite flake tools, although neither of these two studies was of much use in my research.

The three remaining projects did provide me with some useful comparative data and technological insights. For example, Moloney (1988) and Moloney et al (1988), who replicated a series of bifaces using Bunter cobble quartzite from England, contributed valuable information on the knapping behaviour of quartzite, although this was of limited applicability to my work because of morphological differences between our material forms. Ebright (1979, 1987) is so far the only one of these experimenters whose replications were made from quarried bedrock quartzite, similar to the Morrison material from Wyoming. Her work was especially helpful in its treatment of fracture patterns and petrological issues.

Other lithic analysts like Ahler (1989b), Jones (1979), Callahan (1979) and Crabtree (1967) have done some cursory experimentation with quartzite, but only as a small tangent of a large research project. Similarly, Ingbar et al (1989) studied a number of different materials, including two quartzites (one of which was Morrison Quartzite), although they only carried out one biface replication with each raw material type.

It is most unfortunate that researchers have ignored this important raw material, not only because this has restricted replicative studies and essential areas like proper material identification, source location, and basic lithic analysis (Ebright 1987:29), but, more importantly, because it promotes the use of untested analytical methods or procedures which may be generating invalid correlations and conclusions. Pending further evaluation, especially with a variety of raw materials, these methods should be used with caution.

Background and Motivation for the Project

Originally, my thesis research was to be an investigation into lithic scatter formation at a site in the Bighorn Basin region of north-central Wyoming. This area contains a very large outcrop of Morrison formation quartzite (Bies and Romanski 1988; Francis 1983), and was probably an important source area from Palaeoindian to protohistoric times (Prentiss 1992:1). The Morrison geologic formation is of Upper Jurassic age, and is exposed over a large part of the badland areas in the Basin. Both cherts and quartzites commonly occur in different facies of this formation, generally in nodules found easily on the surface, although some minor quarrying did take place (Francis 1983; Frison and Bradley 1980:14), as at the exposed bedrock at the Big Buzz site. The quartzites range from coarse to fine-grained and vary along a scale of colour from grey to rust; all variations in colour and texture can occur within one outcrop or exposure (Francis 1983:41). The most preferred of these materials, however, is the fine-grained, grey quartzite which flakes well and was used in formal and flake tool manufacture as far back as Folsom times (Frison and Bradley 1980). The quartzite used in my experiments was a grey-tan variety of medium to fine grain-size. As noted above, however, there were often variations in texture and colour even within a single piece.

Since none of the available debitage analytical methods had been tested using a coarser-grained material like Morrison quartzite, it was necessary to make sure that there would be no detrimental effects caused by extrapolating the results of chert and obsidian experiments to the interpretation of quartzite debitage. Following an extended search of the available literature on the subject, I came to realize that very little work had been done on this subject and that I would have to carry out my own evaluation of the common approaches to debitage analysis before any future analysis of the Big Buzz quartzite debitage could be accomplished. The new focus of my research would then take into account the different fracture characteristics of quartzite and explore the probable negative effects that this could have on pattern recognition and assemblage interpretation.

CHAPTER TWO:
RESEARCH BACKGROUND AND OBJECTIVES

A/ INTRODUCTION TO LITHIC EXPERIMENTATION

Historical Development of the Experimental Approach in Lithic Analysis

With possibly 99.5% of human history being represented by the Stone Age (Crabtree 1975:105), it is no wonder that researchers have spent so much time and effort trying to understand both stone tools and the debitage produced during lithic reduction in order to interpret behavioural and subsistence patterns. Experiments in the replication of lithic implements can provide specific information about the techniques used to make the items and the possible stages of reduction within such a trajectory. Tool function, resharpening or maintenance, design theory, spatial analysis, typology construction, fracture mechanics, and other issues can also be addressed. "The intended aim of this kind of experiment is the reconstruction of the human and natural agents responsible for the modification, contact traces, damage, wear, and breakage...before, during, and after deposition in the archaeological record. In addition such by-product experiments test the nature and properties of the materials with a view to understanding their potential for human exploitation" (Tringham 1978:182).

Experiments have been of interest to lithic analysts as explanatory devices since the inception of the discipline. In fact, Holmes' work one hundred years ago treated lithic items as historic records that can be used to address questions of time, culture, and the history of peoples (Crabtree 1975:105). Other early signs of interest include the experiments conducted at the Mantes cement factory in 1902 by Marcel Boule and the establishment of a lithic laboratory in Columbus, Ohio in the 1930s (Amick et al 1989:5). Up until the 1940s, lithic experiments were mainly used to distinguish humanly-produced flakes, cores, and tools from naturally formed ones, for example, or to answer questions regarding the identification of cultural traditions, as in the case of certain flakes thought to represent the Acheulian "Clactonian" culture. These flakes were later found to be merely produced by hard hammer percussion during a basic stage in the flaking process, and could not be used to discriminate cultural groups (Ohel 1978).

The next two decades were marked by sporadic, unstructured experiments that tended to be casual rather than scientific in nature. At this time, analysts were still preoccupied with the definition of types for identifying lithic industries, and with carrying out standardized quantitative analyses of measurements and formal attributes which often were not considered for their archaeological significance (Yerkes and Kardulias 1993:90). Debitage studies were practically nonexistent at this time, as the focus remained on finished tools. As modern screening techniques became an accepted part of fieldwork, more and more debitage was recovered from excavations, and its data potential came to be realized by archaeologists.

Researchers also increasingly began to pay more attention to theory rather than just numbers and trait lists. Ascher's 1961 article, "Experimental Archaeology", was a landmark publication since it was the first statement on the use and aims of imitative experiments. Studies from this era were trying to put more emphasis on the context of lithic artifacts and on the mechanics of stone tool production and use (Yerkes and Kardulias 1993:90). With this shift to a focus on behaviour came studies of activity areas (as indicated by flakes or flake scatters), and recognition of a wider variety of knapping techniques, such as Binford and Quimby's (1963) study of bipolar technology. While they did not really understand the conditions that produced these types of flakes, their work was important since most archaeologists were still only identifying conchoidal fracture and percussion methods of flaking. This overemphasis had been prominent in archaeology ever since people began incorporating observations of the manufacture of gunflints at Brandon, England into archaeological interpretations. "Anything that did not have characteristic conchoidal features was relegated to the status of 'chips, spalls, or splinters' and effectively ignored in artifact descriptions and analyses" (Cotterell and Kamminga 1987:681).

By the late 1960s, the work of flintknappers like Crabtree and Bordes began to revitalize interest in the experimental approach, through the training of students in the basics of lithic reduction, the publication of numerous articles detailing the approach, and the production of documentary films on the subject. This opened up many new possibilities for the experimental use of stone (Young and Bonnichsen 1984:135), and helped to develop debitage analysis into a more scientific endeavour.

Another field in lithic analysis that began to develop at this time was "fracture mechanics". The principal proponents of this type of experimental research were specialists in engineering and materials science. By studying the properties of brittle solids, they demonstrated to archaeologists that the basic principles of fracture mechanics could be used to their advantage in understanding the processes of stone tool manufacture and flake formation. For example, observing the type of fracture initiation can tell you about the kind of percussor used, or measuring the crack velocity registered by Wallner Lines can distinguish pressure from percussion flaking (Cotterell and Kamminga 1987:703).

These experiments range from highly controlled ones that are often criticized for not being applicable to most archaeological scenarios, to the more useful ones which can, for example, help to determine which flake attributes or variables to measure (a problem that is pervasive in debitage analysis where no one seems certain about the connection between many variables and the flaking phenomena under scrutiny).

Both fracture mechanics and the more practical types of experiments can potentially offer many contributions to studies of a flake's morphology and its relationship to the actual knapping process. Any knowledge that is gained from these lithic experiments can help to clarify the *organizational* conditions behind such patterning (Amick et al 1989:9). Nevertheless, experimental archaeology has often been criticized for its failure to produce general principles (Pokotylo and Hanks 1989:50) in the same way that other sciences have. This, however, may be related either to a lack of adequate funding and research, or perhaps to the nature of archaeological and anthropological research, which study the material expressions of human groups governed by social, political, or cultural factors for which there are no set rules or standards of behaviour. This problem is especially relevant to the analysis of stone tool manufacture which combines avenues of research from both the "hard" sciences and the humanities. Such an interdisciplinary approach is bound to have conflicting perspectives among the many different proponents, each of whom relies on their own unique kind of research paradigm.

In the 1970s, with the advent of the "New Archaeology", the focus in lithic analysis was becoming more and more like that of the hard sciences. Attention switched to technological issues like determining the reduction technique or technological behaviour that produced certain types of flakes, or trying to define separate stages of lithic reduction (see Collins 1975). In order to make this research more "scientific", people began to develop principles to describe the changes in material remains. The "Frison Effect" , for example, argues that tools change shape as they undergo use and modification (such as resharpening) and that, by the time they finally enter the archaeological record, it will be impossible to tell what these morphological variations were because the evidence is lost through each subsequent removal of material (Dibble 1987:36). In this way, the functional and morphological "history" of the artifact is lost to us.

"Crabtree's Law", a corollary to the Frison Effect, is another principle that helped to advance the status of debitage analysis. This law states that, since the finishing stage in tool production erases the previous step in the reduction process, it is necessary to describe and analyze quarry and workshop debris to discover the technological processes which took place. Debitage is also an important analytical tool because, even though lithic artifacts could be identical in appearance, they could have been made by entirely different techniques (Ritchie and Gould 1985:35-36).

In other words, the only way that we can trace the changes throughout a tool's use-life is by studying the by-products, or debitage, of manufacture and modification. Because of this fact, the importance of waste products as permanent records of technology and prehistoric behaviour began to be appreciated; no longer was debitage relegated to the category of mere waste material that barely deserved mention in archaeological reports. Crabtree (1967:22) hoped that people would view lithic flakes as something with more meaning than just "scraps", and his prediction that these discards would someday help to complete the picture of the past has slowly come true.

As a result of understanding the shortcomings of lithic analyses which *ignored* debitage, experimental (or, actualistic) studies came to be relied upon by growing numbers of archaeologists who found them to be useful methods for explaining variability in lithic debitage assemblages. One of the ways they work best is by identifying technological "constants" which serve as reliable indicators or attributes which can identify a process or behaviour, time and time again. For example, platform lipping is a constant feature that indicates bending fractures associated with soft hammer or billet flaking (Hayden and Hutchings 1989). These types of "fingerprints" are established through replicative experimentation and are valuable in identifying various morphological and technological patterns in lithic debitage assemblages.

By understanding such variability and the patterns that it takes, it is then possible to decipher the technological processes behind it. This, in turn, allows researchers to relate archaeological patterns to broader questions regarding cultural organization, thus linking method and theory in a controlled manner (Prentiss and Romanski 1989; Amick et al 1989; Ingbar et al 1989). Since we have no means of directly recovering prehistoric human behaviour, replicative experiments offer the best possible way to approximate the past, by providing interpretive models and reliable means of inference (Ahler 1989b; Magne 1989). In fact, Luedtke (1992:80) argues that "replication now plays an irreplaceable role in determining how stone tools were made, understanding site formation processes,...and generally maximizing the cultural information we derive from lithic remains".

Do Actualistic Studies Assume Too Much?

Despite the fact that experimental studies have been able to offer unique contributions and a more dynamic view of the past, particularly when used in combination with refitting or microwear analyses (Yerkes and Kardulias 1993:89), the progress over the last ten years or so has still come under attack on a number of different issues. Thomas (1986:247-251), in particular, has condemned replicative experimentation as outdated, normative and irrelevant. While not all would necessarily go so far in their criticism, potential problem areas must be acknowledged.

As with all experiments, there is the ever-present possibility of researcher bias, especially in terms of cultural subjectivity. In lithic studies we are dealing for the most part with the technological base of extinct cultures, and also suffer from the absence of useful ethnographic data on lithic technological organization (Pokotylo and Hanks 1989:49). Consequently, any models or inferences we make could be biased by our own preconceptions or backgrounds. Arguments have been made that the influences of our modern-day cultural system can unduly affect the way we interpret the factors determining prehistoric assemblage compositions. Furthermore, experimental knappers can be affected by their personal knowledge of how artifacts are thought to have been made and, conversely, they will interpret archaeological remains in terms of their own production code, without clearly describing their interpretive models (Young and Bonnichsen 1984:135).

This concern is echoed by Luedtke (1992:80), who noted that the work of most knappers becomes largely subconscious, thus making it difficult for them explain what they are doing. There will also be the problem of different researchers perceiving and interpreting the same experimental data in different ways. Since our perceptions of the evidence can vary, it is necessary to remember that by comparing experimental results with archaeological collections, we are simply trying to develop homologs, not analogs, for explaining the past (Amick et al 1989:8).

Further problems arise when we examine the experimental design of lithic replicative studies. One of the main criticisms here is that there is no room allowed for flexibility in these highly-controlled and very goal-oriented experiments, and there seems to be no consideration for the effects of unplanned influences or situational factors (Magne 1989:16; Amick et al 1989:7; Mauldin and Amick 1989:67). Stoneworkers are often required to change their reduction strategies or production goals in mid-stream, for example, even though no lithic experiment has allowed for this type of reorientation before. This has to do with the fact that researchers must strictly control the variables under investigation in order to minimize the complexity of the situation and thus determine the effects of the variables on tool manufacture and its by-products. As a result, there is a certain amount of artificiality inherent in the method and results of most controlled experiments, such as Bonnichsen's (1977) "Stainless Steel Indian" study.

Similar machine testing was later carried out by such researchers as Speth (1975,1981), Dibble and Whittaker (1981) and Dibble (1985). Their fracture mechanics experiments involved dropping steel balls from electromagnets onto a glass plate or core in order to control and relate as many variables as possible. As mentioned earlier, while they may have succeeded at clarifying the relationships among some variables in the laboratory, their work is often criticized as being too remote from the actual conditions surrounding tool manufacture in the past (Luedtke 1992:81). This reflects, nevertheless, the increasing concern with deriving technological information from knapping by-products.

Are actualistic studies of any use to archaeologists, then, and do they accurately reflect the relationships that exist in the archaeological record? This is an example of some of the concerns that have been voiced against replicative experiments. Ingbar et al (1989:118), for example, saw many problems in the conceptual and operational aspects of most debitage analyses, and argued that experimentally-derived models may only be applicable to other controlled cases. This might be due to the fact that there are factors in archaeological assemblages that cannot be controlled for, as there are in lithic experiments. Making the inferential leap from the latter to the former may therefore be jeopardizing the validity of our claims. Other researchers have voiced their concerns (eg. Speth 1981:16; Patterson and Sollberger 1978:103) over this issue, stating that the results of debitage studies have limited use for application to actual archaeological sites or complex artifact morphologies. This is particularly true in relation to the problem of over-controlled experimental results (often derived from a single knapping event) being unable to interpret the multicomponent, mixed assemblages found in the archaeological record. In addition, replicative experiments do not take the problem of recycling into account, even though archaeologists often find evidence that some flakes were used in multiple contexts or were selectively removed from manufacturing debris for tool use.

Reliability and Validity

It is encouraging that these types of issues are slowly being taken into consideration, as archaeologists continue to question the methods and assumptions they use. It is unfortunate, though, that most researchers are simply not aware of issues like reliability and validity. According to Nance (1987; Nance and Ball 1986), these concepts originate in the context of measurement theory and are used to ensure that experiments are well-designed and theoretically sound, and that any conclusions that are drawn will go toward developing accurate, consistent, and precise statements about our data.

A measurement or procedure is considered to be ***reliable*** only if it consistently gives the same results when applied to the same object or data. Whereas *random error* (that is, all the chance factors involved in making measurement mistakes) is one factor that can cause the results to be inconsistent or unreliable from one time to the next, *non-random error* (such as interobserver bias or changes in an analyst's perceptions over time) is a factor that affects an instrument or procedure's ***validity***, or, its ability to measure the phenomenon it claims to. For example, if a measurement such as an entrance exam score is highly correlated with a student's Grade Point Average, the exam is considered to be a valid predictor or measure of academic success (Nance 1987:284-285). Similarly, in lithic analysis, if a measurement like an estimate of the amount of cortex on a flake is highly correlated with a certain stage in the lithic reduction sequence, then the percentage of cortex cover is argued to be a valid indicator of manufacturing stage. Therefore, precise meanings of concepts and standards for comparison are extremely important in the establishment of instrument validity (Amick et al 1989:3).

It is unfortunate that so little attention has been paid to such a fundamental aspect of experimental design, although this may be overcome in time, "given the nascent condition of debitage studies as a whole" (Ingbar et al 1989:118). At present, lithic experiments are the best means that we have for providing insight into: prehistoric lithic reduction trajectories; the activities and distinctive debitage associated with stages of manufacture; and defining and evaluating which variables are significant in lithic analysis. While controlled experiments may not provide us with all the answers we seek, they are still the most effective way of establishing a baseline of lithic artifact variability (Yerkes and Kardulias 1993:93), and are capable of denying or clarifying the relationships that appear to exist in uncontrolled circumstances. As archaeology, and lithic analysis in particular, evolves and matures, the number of valid theoretical constructs will no doubt increase to a more satisfactory and beneficial level.

B/ FUNDAMENTAL THEORETICAL CONSTRUCTS IN DEBITAGE ANALYSIS

Introduction

Debitage, or the residual lithic material resulting from tool manufacture (Crabtree 1972:98), has long been recognized by archaeologists, but it has only been in the last decade or two that researchers have begun to seriously study its variability and connections to lithic reduction processes. Unfortunately, many analysts still find that debitage is among the least studied lithic items, despite its potential to render primary information (eg. Magne 1989:15). Gradually, more and more people are acknowledging it as an important focus of lithic analysis that is capable of informing us about specific issues like reduction strategies and flaking techniques, and more general questions regarding the integration of tool manufacture and use at a landscape level (Amick et al 1989:67). Different occupation zones at multicomponent sites can be isolated by analysing the debitage, and lithic scatters can be differentiated and tied into settlement systems (Yerkes and Kardulias 1993:93). The possible applications of debitage analysis are numerous and intriguing. Debitage can answer questions regarding culture history, trade and exchange, occupational histories, social organization, craft specialization, design theory, and overall organization of technology, to name but a few topics.

In general, debitage is more diagnostic than tool types because "the reduction techniques preserved in debitage are the stable cultural markers of prehistoric knapping behavior" (Flenniken 1985:267). Similarly, the flake retains more diagnostic features than the flake scar because the striking platform usually adheres to the flake. It therefore bears special traits, like the positive bulb of force (which can identify the mode of detachment and the stage of manufacture), and it can designate the area contacted by the fabricator, or allow the inspection of both the dorsal and ventral surfaces (Crabtree 1972:107).

In addition, since debitage is an immediate by-product of the manufacturing process, it is less likely to be curated or transported away from its original knapping location (Magne 1989:15; Odell 1989:163). Although many of the finished products were moved elsewhere, the flakes and rejected items left behind can provide specific information about what the knappers were producing and how their tasks were being accomplished (Shafer 1985:293). Microdebitage, in particular, yields useful information on activity areas and is "a stable and long-lasting signature of past cultural activity imprinted on the sedimentary environment, and is not subject to some biases inherent in macroscopic cultural materials (such as selective collecting, re-use, etc.)" (Fladmark 1982:215).

Flake Typologies and Technological Constants

Classification systems have been used as essential organizational tools in archaeology for decades in order to help identify distinct patterns of behaviour or technology which, in turn, can be used to reconstruct aspects of cultural development and interactions. However, classifications themselves do not necessarily have any relevance to archaeological problems, since they are merely groupings of objects (Hayden 1993:90). Nevertheless, most prehistorians do not agree on the goals of typology, either (Cahen and Van Noten 1971:211). Sometimes it is considered only as an initial descriptive and classificatory stage while, at other times, it is hoped that the typology itself will serve as a way to identify or define prehistoric industries, as was the case in lithic analysis during the 1950s and 1960s (Yerkes and Kardulias 1993:90). Ideally, typologies should be carefully constructed by determining the appropriate artifacts and attributes that can combine to form "types"; these, in turn, should enable archaeologists to answer questions about prehistoric culture (Hayden 1993).

The focus of most typological approaches used in debitage studies today is generally the identification of technological constants (Ingbar et al 1989) which are useful in assessing the technological derivation of debitage assemblages (Ahler 1989:210). These constants can include any flake features (such as bulbs of percussion, cortex cover, or dorsal scar counts) or distributional data which are thought to be reliable indicators of technological actions or intent. At the most fundamental level, then, the goal of lithic reduction experiments should be to construct good typologies that provide reliable measurements or characterizations of flakes, so that we can then use them in modelling assemblage formation (Magne 1989:28).

Unfortunately, the connection between typologies and inferences is often poorly articulated; therefore, we need to design experiments that can "help to identify poorly constructed classificatory devices and potential ambiguities related to the interpretation of these typologies" (Amick et al 1989:1). My research here has been developed in an attempt to improve replicative experiments in such a way. I, too, believe that it is essential to pinpoint the weaknesses in our instruments or procedures, and also in the way they are used to interpret flake collections. For example, it is not enough to be able to identify a piece of debitage as a "biface thinning flake", since assigning this label or this place within a typology does not always explain the range of variability associated with the entire assemblage. We need to develop theories and models of stone tool technology that can reliably identify constants and validly apply them to whole collections of lithic items; until then, considerable doubt will persist regarding the potential of inferring past behaviour from lithic analysis (Pokotylo and Hanks 1989: 49).

Lithic Reduction Sequences and Stages of Manufacture

In an effort to compensate for this atheoretical state, and to improve our understanding of lithic assemblages, analysts began to focus on how stone tool technology was organized and ordered at the level of settlement systems. Flow models have been found to provide useful means of characterizing the technological processes of lithic industries, from the acquisition of the raw material, to the thinning and shaping of a tool or the preparation of a core, to the discard of the finished product (Newcomer 1971; Bradley 1974; Collins 1975; Schiffer 1976; Magne 1985; Magne and Pokotylo 1981; Rankama 1990:115). Flow charts, which model lithic reduction processes such as these, are linear in nature, with one step "logically" leading to the next; for example, tool preforms undergo hard percussor "roughing out", followed by secondary thinning or soft hammer refinement, and finally they are prepared for hafting, serration, aesthetic modifications and so on, usually by pressure flaking. The size of waste flakes is argued by some to decrease proportionally as this reduction sequence proceeds (Stahle and Dunn 1982).

Researchers admit that they are operating under a limited and generalized reduction model, although this is perhaps the only way in which to understand the various discrete "stages" of lithic activities (Magne 1985). These stages serve as analytical categories which are thought to reflect the changing objectives and techniques involved in lithic reduction (Stahle and Dunn 1982:84), as well as the position of waste flakes in the reduction sequence, although they are admittedly arbitrary experimental designations (Raab et al 1979).

Stage Typologies and Debitage Categories

Despite the potential problems associated with defining manufacturing stages, much time and effort has been spent (using this as a theoretical guide) trying to discern the technological origins of two "separate" types of debitage: tool and non-tool waste flakes.

Tool manufacturing debitage is generally defined on the basis of several morphological attributes such as striking platform features, size, thickness, curvature, flake scar patterns, and so on (Sullivan and Rozen 1985:757). In addition, certain flake types (such as biface thinning flakes, billet flakes, or pressure flakes) are used to identify specific knapping activities that occurred at some point along the proposed manufacturing trajectory. These flake types are often subjectively identified, and their definitions tend to be different from one researcher to the next, making comparisons between experimental data sets quite difficult or even impossible (Raab et al 1979:171).

Non-tool debitage, on the other hand, is generally categorized according to a single attribute: the amount of cortex cover present on a flake's dorsal surface (Sullivan and Rozen 1985:756). Typologies for non-tool waste, or core reduction flakes, classify the debitage into "primary", "secondary" or "tertiary" categories, in order of decreasing percentages of cortex (Frison and Bradley 1980; Boksenbaum 1980; Francis 1983; Stevenson et al 1984; Tomka 1989), although occasionally a residual category for "shatter" is included. These categories are said to represent a very stable sequence of flake removals, whereby, in one nomenclature system, primary flakes are assumed to be the first ones removed from a core or blank (they therefore have the greatest amount of cortex), secondary flakes with less cortex are the next to come off, and finally tertiary (noncortical) debris is the last in the sequence to be removed. While, no doubt, there are exceptions to this particular typological approach, the dominant theoretical construct states that the amount of cortical cover acts as the most reliable indicator (or constant) of the stage in the reduction sequence at which the flake was removed, in a similar manner to the way that decreasing flake size has been correlated with progressively later stages of reduction.

C/ PROBLEMS WITH DEBITAGE ANALYSIS IN GENERAL

Virtually all the procedures, typologies, and approaches used in debitage analysis have come under a great deal of attack on a number of specific issues, ranging from terminological difficulties and other problems that threaten their reliability, to criticism of the very theoretical constructs that guide their logic and ensure valid inferences. A summary of these weaknesses is provided below.

Particularistic Nature of Lithic Experiments

As mentioned previously, one general problem with methods of debitage analysis is regarding their narrow range of applicability. Unfortunately, most experiments tend to produce results that can only be applied to similar experimental data. Consequently, the lack of understanding of the range of variability in flaking debris, either in experimental settings or in the actual archaeological record, can lead to inaccurate interpretations of lithic assemblage formation (Magne 1989:16).

Researchers need to find a way to make their experiments more flexible and realistic, by taking into account the fact that situational factors are often just as responsible for assemblage composition as mechanical or technological factors. Until we can learn to consider situational conditions such as mid-stride changes in reduction technique or reorientation of production goals during reduction , as well as production rejects and failures, varying levels of stoneworking skill, or the sequence of flaking tool use (Magne 1989:16; Amick et al 1989:7; Ahler 1989b; Mauldin and Amick 1989: 67), we will never be able to accurately account for variability in debitage assemblages.

Terminological Comparability Between Researchers

Prior testing and attempts at applying various methods of debitage analysis have revealed a number of inconsistencies and unreliable aspects which may lead to inaccurate interpretations of flaking debris assemblages.

Even on a very basic level, such as that of the terminology used in this type of research, the work is often unstandardized, subjective, and inconsistent (Luedtke 1992:80). While this may at first seem like a trivial criticism, it is important to remember that typologies must be compatible and comparable in order to allow meaningful inferences and comparisons to be drawn.

Unfortunately, some researchers will not provide any definition of the types or attributes that were used in their studies (eg. Boksenbaum 1980; Stahle and Dunn 1982:85). To illustrate the dangers of this terminological dilemma, consider a few common labels used in debitage studies. The flake category of "shatter" , for example, has also been referred to by various analysts as "debris", "waste material", "residual items","angular waste" (Minor and Toepel 1982:9), "nonorientable flakes" (Prentiss 1988, 1993), "chunks" (Tomka 1989), "dregs" (Boksenbaum 1980), and "LILFs", or "limited information lithic fragments" (Shelley 1990:191).

This confusing array of terms can easily cause problems when researchers wish to compare their data with someone else's, and also when the definition of a particular flake type is different from one study to the next. For example, Magne (1985; Magne and Pokotylo 1981) uses "shatter" to indicate all flakes without striking platforms, as opposed to Sullivan and Rozen (1985) who would call such items either "flake fragments" (or "medial-distal fragments", according to the modified Sullivan and Rozen typology [Prentiss 1993]) or else nonorientable "debris". Similarly, Newcomer and Karlin (1987:33) include this kind of flake under the category "chips", which can also cover tiny complete flakes, shatter, and fragments of larger whole flakes.

If no one can even agree on what to call the various kinds of lithic debitage, then we will never be able to compare and contrast our data sets. Fish (1978) and Beck and Jones (1987) recognize the dangers caused by this type of bias, and warn people that they must allow for classificatory divergence between analysts, especially whenever comparing published data to their own. The resulting problems with classification and communication will only prevent debitage analysis from becoming a reliable and valid interpretive tool.

Bias, or systematic error, which affects the validity of an approach, can also arise from differing understandings of the definitions or attributes being examined, both between observers and with regard to an individual's changing perceptions over time. Such differences in the application of debitage classifications can "introduce serious levels of systematic error that potentially can obscure all but the most robust archaeological patterns" (Beck and Jones 1987:259). The work of researchers like Ahler (1989b) is notable since he provided explicit definitions of shatter, bipolar flakes, biface thinning flakes and other classes of debitage. Brown et al (1982) took this even further by not only providing definitions of their types, but also by including a section which indicated potential sources of error related to each classification.

It is apparent, then, that most studies have far to go in the standardization of debitage terminology; nevertheless, this is one issue that must be reconciled in order for lithic analysts to generate unbiased, replicable results and interpretations of their data.

Flake Selection Criteria

To further complicate matters, differences in the choice of flake types used in experimental debitage analysis can also introduce bias or systematic error if researchers compare their data and conclusions to those from other studies which used different variables and flake attributes.

For example, some researchers have chosen to use all types of flakes, either whole or fragmentary, when typologically or quantitatively examining debitage collections (see Boksenbaum 1980; Jelinek et al 1971; Mauldin and Amick 1989; Stahle and Dunn 1982; Sullivan and Rozen 1985; Prentiss and Romanski 1989; Prentiss 1992; Ahler 1989).

On the other hand, some experimenters select only complete flakes, that is, flakes with a striking platform, a bulb of percussion, and (usually) a feather termination (eg. Hiscock 1986; Odell 1989; Stahle and Dunn 1982; Raab et al 1979), despite the fact that they are ignoring a great deal of variability by biasing their sample in this way. Similarly, Magne (1985; Magne and Pokotylo 1981) and others (eg. Holm 1990; Stevenson et al 1984) choose to select only flakes with striking platforms, and they argue that the termination is not a source of valuable information on reduction stages. To counteract the negative effects of this type of selection bias, Hayden and Hutchings (1989:254) strongly recommend that analysts employ more realistic polythetic approaches, since "details of initiation areas can often be obscured by eraillure scars, shattering, breakage, and other common observational impediments".

To illustrate how these differences in selection criteria can be problematic, consider Ingbar et al 's (1989) experimental work in which they developed a regression equation to statistically explain debitage variability, based on platform-bearing flakes only, and then proceeded to apply it to the entire assemblage, including all types of flakes. Such inconsistencies can lead to misleading correlations between the attributes under study as well as between a variety of experimental databases, each derived from a different research design.

Lack of Consensus Regarding Technological Constants

Another pervasive problem in debitage analysis is the fact that there is no consistency or agreement about which attributes are the most useful for identifying reduction stages or knapping techniques. Quite often attributes are selected for analysis without any explanation about what each of them contributes to the interpretation of the lithic collection (eg. Holm 1990).

The most prominent aspect of this problem is the derivation of many variables or attributes thought to reflect lithic production behaviour; "Some of the attributes are based on experimental evidence, some are based on logical arguments, and some are based on intuition" (Mauldin and Amick 1989:67). Even when experiments *are* used, they are often derived from a very small number of reduction events or from small flake samples (eg. Tomka 1989; Magne 1985; Ingbar et al 1989; Odell 1989). Furthermore, the types of reduction strategies are generally limited to either biface manufacture or core reduction, and therefore the results can not be reliably applied to other situations such as bipolar or blade-core debitage, for example (see Magne 1989:19).

Many of the earlier studies tried to isolate single variables to serve as constants or predictors (eg. Magne 1985), although it is now generally agreed that *combinations* of attributes provide the most reliable method for identifying knapping behaviours, from core reduction (Mauldin and Amick 1989; Ingbar et al 1989; Odell 1989), to soft hammer percussion (Hayden and Hutchings 1989), to bipolar reduction (Hayden 1980), and finally to pressure flaking (Tomka 1989).

Problems arise, however, when particular attributes must be selected from a long list of possible choices. While most researchers do not want to choose variables which are redundant or non-exclusive, there are those who blindly rely on "standard" measurements and attributes such as weight, thickness, and a variety of platform measurements (often studying as many variables as possible in the hopes of deriving meaningful patterns

[Magne 1985:26]), even though they cannot provide reliable or valid explanations of debitage variability. This issue is especially relevant to those who analyse very large debitage assemblages which sometimes number in the thousands of flakes, because it takes a tremendous amount of time to measure and record these attributes.

Nevertheless, even with the short-listed attributes, there is seldom any consensus among lithic analysts as to which are the most useful. People disagree about whether or not to measure platform angles (eg. Dibble and Whittaker [1981] vs. Magne and Pokotylo [1981]), dorsal scars (eg. Mauldin and Amick [1989] vs. Magne [1985]), and many other attributes. However, it is hoped that further research and experimentation will identify the most reliable and discriminating variables and attributes that can provide insight into the technological aspects of flake production and manufacturing sequences (Speth 1972:57; Magne 1989:18; Odell 1989:170).

Explanations of Cortical Variation

One of the most fundamental theories in debitage analysis is that it is possible to identify stages of lithic reduction by classifying flakes according to how much cortex cover they have. The underlying assumption is that "cortex in any amount is overwhelmingly present in early or core reduction stages, and only rarely in other stages"(Magne 1989:17). However, many different factors can influence or directly determine the amount of cortex.

To begin with, the original morphology of a core or blank is an important determinant of cortex cover (Dibble 1985; Odell 1989; Beck and Jones 1990:293; Ebright 1987; Roe 1988; Prentiss and Romanski 1989; Ingbar et al 1989). This is because different core materials and sizes have different amounts of cortical rind and, in some cases, there may never have been *any* cortex there at all. For example, if the knapper is reducing a quarried chunk of bedrock that possesses no cortex, then typologies or approaches that rely on measuring cortex cover will be unable to provide a valid characterization of the flake assemblage and its technological origins. Similarly, the percentages of primary and secondary flakes will be higher for the reduction of a cobble than they will be for a flake blank derived from a prepared core.

Furthermore, cortex can be present at any stage in the reduction "sequence"; it is not always completely removed prior to soft hammer or pressure flaking refinement (which is generally thought to be carried out once the blank is decorticated). In fact, experiments have demonstrated that this variable is only a reliable indicator of the earliest stages of reduction (Mauldin and Amick 1989; Odell 1989) and with large initial core materials.

Despite the many problems associated with debitage analysis, it is important to keep in mind that this field is responsible for generating many useful analogies and comparative studies. As Magne (1989:27) put it, "...models are not meant to be reality, but are meant to have reality bounced against them".

D/ METHODOLOGICAL SCHOOLS IN DEBITAGE ANALYSIS

From the mid-1970s on, controlled experiments became the norm. Most notable among the early ones is the "Mass Analysis" approach developed by Ahler in 1972 (see 1989b:93-98), and other size-grading approaches (eg. Henry et al 1976; Patterson and Sollberger 1978; Stahle and Dunn 1982). These are unique among debitage analysis methods because they examine the characteristics of the assemblage as a whole, rather than studying the features of each individual flake. The basic premise of these approaches is that, when a debitage assemblage is screened through different sizes of mesh, it will become sorted into groups (based on flake size) that reflect the various reduction stages represented in the assemblage. These patterns can then be used to distinguish one type of reduction from another based on the composition of the size grades. While such methods do not explore small-scale variability and may allow fewer technological inferences to be made (since they indicate *general* trends in the data), they are extremely efficient in terms of time, effort, and money. Those who have employed mass analysis in their debitage studies have found that the objective, reliable results outweigh the minor disadvantages (eg. Stahle and Dunn 1982).

The search for regular patterning in debitage assemblages gained popularity, and many lithics researchers concentrated their efforts on finding technological constants (Ingbar et al 1989:119), or experimentally-generated "signature assemblages" thought to identify various knapping processes (Tomka 1989:157).

One method of debitage analysis that relies on technological constants is individual flake attribute analysis. This method looks at the systemic context of individual flakes within lithic reduction sequences. By examining flake morphology and scar counts, Magne (1985,1989) and others (Flenniken 1981; Magne & Pokotylo 1981) use constants such as cortex cover or striking platform characteristics to classify flakes according to their sequential position in a reduction trajectory. This involves the assumption that lithic reduction occurs in discrete stages along a predetermined knapping sequence. They argue that each flake carries significant information on such knapping behaviour, and that debitage is technique-specific. As a result, they claim that this method is successful at indicating how each flake was detached, and at what point in the reduction sequence this detachment occurred. Nevertheless, there are some drawbacks to this approach. Foremost is the probability that their samples are biased, since they sometimes use complete or platform-bearing flakes and ignore all others. Taking all of the measurements on each flake can also be extremely time-consuming (especially since debitage assemblages tend to be very large) and redundant. These concerns are particularly relevant when some researchers even doubt the validity of the stage concept itself, although I will discuss these issues further in Chapter Four.

Based on similar dissatisfaction with stage approaches, Sullivan and Rozen (1985) developed their own method based on archaeological assemblages they had been studying. They argue that it is very reliable and less time-consuming than stage methods, although you will get similar results. Instead of focusing on cortex cover or scar counts, they simply looked at the degree of completeness of each flake. Researchers following this approach would then use a hierarchical classification scheme to sort all flakes into mutually-exclusive categories. This "interpretation-free" method was thought to be successful at identifying the technological origins of "distinctive assemblages" of debitage (Sullivan and Rozen 1985:755), especially distinguishing core reduction from bifacial reduction, although later testing through experimental replications found that it may not actually measure what it claims to. The fundamental flaw seems to be the lack of experimental verification of their claims; archaeological data should be used as a control (Yerkes and Kardulias 1993:93), not as the unquestioned basis for an analytical technique. Prentiss (1993; Prentiss and Romanski 1989) experimentally challenged the reliability and validity of the Sullivan and Rozen technique and found that, while it could produce the same results each time it was used , it did not accurately identify the technological actions it claimed to. By dividing each of the flake categories into four size grades, Prentiss then found that he could improve the validity of this approach, and he dubbed the new version the "Modified Sullivan and Rozen Typology" , or MSRT.

His work has contributed a great deal to the direction of debitage analysis, by demonstrating the benefits of testing both the techniques or approaches that we employ, as well as the theoretical assumptions on which we base our methods and conclusions. Hopefully this emphasis on reliability and validity testing will become incorporated into the research design of more and more experimental projects, as it has with my research. In this way, we will be ensuring that the interpretations or principles we develop and promote are truly accurate and precise characterizations of the data and the relationship of the variables under study. By locating sources of information that can be used to validate our analytical approaches, such as information gained from research in other sciences or from within archaeology itself (as in the case of replicative experiments), we will be able to use our measuring instruments and analytical procedures with a great deal more confidence.

E/ THE IMPORTANCE OF RAW MATERIAL CHARACTERISTICS IN DEBITAGE ANALYSIS

One avenue that was recommended for future research was to explore the effects of raw material variability on the validity or success of these debitage analysis methods (Prentiss 1992; Magne 1989; Ahler 1989a, 1989b). It is interesting to note that this issue has essentially been ignored in *contemporary* lithic analysis, even though the early experimenters did recognize its importance. Crabtree (1975:108), for example, acknowledged the fact that, while there are many factors to be considered when determining the technique of lithic manufacture, we "must first evaluate the vast differences in lithic materials. I cannot stress this too much". Even earlier, Goodman (1944: 416) had more explicitly called for the investigation of the physical properties of lithic materials to be regarded as a significant sector of archaeological research.

Unfortunately, in most cases, any consideration that is given to the subject of raw material seems to be in relation to its availability (for example, see Kuhn 1991; Vehik 1984; Francis 1983; Magne 1989), or to patterns of lithic procurement and group interaction (see Ellis and Lothrop 1989; Luedtke 1992:76; Toll 1978:50). Little attention has been devoted to exploring the petrographic nature of material types, or the direct influence that this has, not only on flaking quality, design theory, or assemblage content, but subsequently on the validity of the inferences generated by the various methods used in the analysis of debitage. Studies such as that by Ingbar et al (1989), for example, show an interest in variation in raw material treatment from the same and different sources, yet they mention nothing about the qualities or behaviour of the materials, two factors which can affect the basis of experimentally-derived models like the ones they use. Even when archaeologists do acknowledge that flaking characteristics or qualities (such as hardness, brittleness, texture, homogeneity and so on) must be evaluated in order to explain variation in our assemblages, most do not follow up on the idea (see Francis 1983:145; Dibble 1987:36). The result has been the establishment of misleading and often invalid techniques for analysing debitage. If we are to provide both reliable and valid approaches to the study of lithic materials, then we must begin to evaluate all aspects of debitage analysis, including the characteristics of *all* stone types, not just those which flake easily or show fracture features clearly.

Most of what we do know about lithic material behaviour comes from the field of fracture mechanics. Cotterell and Kamminga (1987:677), for example, inform us that materials with the desired properties of homogeneity and isotropy were the most favoured, since they are less direction-dependent and therefore fracture in a predictable manner. When prehistoric people were able to select materials with these ideal properties, they did, but many times they were limited to inferior sources (Collins 1975:19). Sharpness and flakability were found particularly in natural glasses like obsidian (even though it was one of the scarcer lithic materials), and cherts, which have greater fracture toughness than glasses. Most experimental work has involved these stone types (Kamminga 1982:22; Runnels 1985:100), even though tough, siliceous materials such as quartzite were also commonly used in prehistoric tool making; consequently, replicative analysts should be cautious about relying on chert and obsidian in their experiments, when the vast range of lithic materials employed prehistorically is considered.

The importance of these other rock types has been highly underrated (Strauss 1989:25). In Australia, for example, Kamminga (1982:22) noted that any lithic type above a 5 on the Moh's scale of hardness could have effectively accomplished the tasks at hand;

however, as yet, "no experimental design has included the testing of a broad range of stone types". Even Ahler's on-going experiments over the last twenty years have failed to expand enough to give serious consideration to materials other than cherts and "flints", despite the fact that "the [archaeological] target samples reflected a broad array of techniques applied to more than a dozen kinds of local and nonlocal raw materials" (Ahler 1989: 200).

Prehistoric Use of Quartzite

One of the most misleading assumptions in lithic analysis is that there was always a preponderance of flint, chert, and obsidian in European and North American stone industries (Kamminga 1982:22). Obsidian, in fact, is relatively rare on a worldwide basis (Taylor 1976:vii). Roe (1988:2) has similarly argued that "if flint had been essential to tool manufacture, the Paleolithic would never have got off the ground in Plio-Pleistocene Africa; indeed, ...one could argue that half of the whole timespan of the Paleolithic had elapsed before any toolmaker first picked up a nodule of flint, as opposed to other siliceous rock".

Raw material is, in fact, a very broad category that can refer to sedimentary, igneous, or metamorphic rocks, as well as minerals, glasses and crystalline aggregates (Collins 1975:19). It is surprising, then, how narrow our view of prehistoric lithic exploitation has become, particularly over the last two decades. Quartzite, in fact, has played an extremely important role in ancient subsistence systems (Crabtree 1967:11) and constitutes one of the most widely used aboriginal lithic materials (Ebright 1987:29). Quartzite tools are known from the earliest hominid levels and from every inhabited continent (Toll 1978:49). Semenov (1964:34), in fact, has argued that people had little choice in using this material.

In the "Old World", quartzite has been described as the most widespread raw material used during the early Paleolithic, in particular, and has been reported in

Czechoslovakia, Hungary, the English Midlands, southern Russia, Olduvai Gorge, Sterkfontein, Swartkrans, and in India, Asia, and Peru (Cummins 1983:199). Moloney et al (1988:25) have also described its presence at all Acheulian sites in England, France, Spain and Portugal, including Terra Amata, Lazaret, and Ambrona. In the Early and Middle Neolithic, it was an important raw material in Greek Macedonia, as well (Watson 1983:123).

It is clear that the choice of stone tool material is limited by local geology. Whenever and wherever high-quality, siliceous materials like flint, chalcedony or obsidian were available, they were understandably exploited. However, in regions where flint or obsidian were scarce or absent, quartzite appears to have been heavily relied upon, as was the case in Africa (Inskeep 1988:236-237; Brothwell 1983; Strauss 1980:69), Australia (Kamminga 1982:22), Britain (MacRae 1988b:98; Moloney 1988), and Northern Europe (Coulson and Skar 1990).

As a result of the general research bias toward flint in Old World assemblages, many researchers noted the difficulty in recognizing and analysing artifacts made of quartzite or similar rocks (Wymer 1988:11). This bias developed to the point where countries such as Finland, Norway and Britain had, until recently, been "relegated to a marginal role in Stone Age research because of their lack of flint-dominated assemblages" (Coulson and Skar 1990:71). In an effort to resolve this dilemma, the "Flint Alternatives Workshops" were set up. Their goal was to help archaeologists who worked in areas where flint comprised only a very small percentage of archaeological finds. They have focused not only on quartzite, but also on quartz, porphyrite, rhyolite, slate and other materials.

Unfortunately, nothing similar to these workshops has ever been encouraged in the North America, where quartzite was also a heavily-utilized lithic resource from Palaeo-Indian times, on the High Plains, for example (Frison 1982), during the Archaic (especially in the Eastern Woodlands) (Ebright 1987:29), through to the historic era (Ebright 1987;

Toll 1978). It was exploited a great deal in the Rocky Mountain regions in Colorado (Toll 1978:50) and southeastern Alberta (Bonnichsen 1977), as well as in New England (Strauss 1989), and at the Sheguiandah quartzite quarry in Ontario (Lee 1954, 1955).

Although quartzite was found all over both North and South America, archaeologists tend to give this material only cursory attention, most often just relegating it to a brief mention in their reports or appendices. Rock types get lumped into lithomechanically equivalent groups without explanations or definitions of what the analyst means by the ambiguous categories (Shelley 1990:188: Toll 1978:50). The ubiquitous "Other Materials" category (see Boksenbaum 1980, for example) is relied on far too heavily in lithic analyses, and serves to reinforce the idea that quartzite and other similar rock types are trivial materials with no importance of their own. When researchers mention non-obsidian or non-chert debitage but won't specify what kind of material it actually is, this only serves to further misguide our interpretations of the past. If only the "nice" raw materials like chert and obsidian are given consideration, both in site reports *and* in replicative experiments, then quartzite and other similar stone types are doomed to remain invisible not only to lithics researchers, but to anyone interested in developing an accurate picture of ancient subsistence and technological organization.

In terms of lithic experimentation, the bias caused by focusing on chert and obsidian has far-reaching negative effects. Most of the important replicative studies, unfortunately, use only these two kinds of materials. This is true both of the earlier lithic experiments carried out in the 1970s and early 1980s (for example, see Behm and Faulkner 1974; Chandler and Ware 1976; Raab et al 1979; Patterson and Sollberger 1978; Gunn 1975; Henry et al 1976; Sieveking 1980; Dibble and Whittaker 1981), and those, too numerous to mention here, which were carried out more recently. To illustrate this continuing trend, consider the fact that, out of all the replicative experiments in Amick and Mauldin's (1989) landmark volume *Experiments in Lithic Technology* (eg., Odell 1989; Hayden and

Hutchings 1989; Mauldin and Amick 1989; Prentiss and Romanski 1989; Baumler and Downum 1989; Ahler 1989; and Tomka 1989), only one (Ingbar et al 1989) incorporated a material other than chert or obsidian (in this case, quartzite).

Raw material type, however, has a tremendous influence on the lithic reduction process, a fact that has long been recognized (see Pond 1930, Semenov 1964; Crabtree 1967; Ranere 1975). These oversights and biases found in more recent lithic analysis, then, are somewhat surprising, since one would expect that a research field would *expand* its horizons, not increasingly shrink them to the point of abstraction. Variability caused by the different physical properties of raw materials has often been *suggested* as an area in need of further exploration (see Henry et al 1976:60; Strauss 1980:71; Runnels 1985; Magne 1985, 1989:27; Prentiss 1993; Ingbar et al 1989: 97) , nevertheless, studies using non-chert and non-obsidian lithic types are still quite rare. As rock types other than chert or obsidian are slowly and gradually gaining recognition as valid materials for prehistoric stone tool manufacture, the influence of their composition and flaking behaviour on technology and, thus, assemblage variability, is only just beginning to be tested by experimentation (Moloney et al 1988:25).

By incorporating information from outside the discipline as well as from within experimental lithic analysis, we are making use of a variety of criteria with which to determine the validity of claims made by debitage analysts (after Nance 1987). We need to use *internal criteria* generated by replicative experiments to re-evaluate how raw material qualities affect the form of a stone tool, its use, and the retouch used to refine it (Jones 1979:835; Brown et al 1982:5), and also to see how raw material is related to general tool manufacturing processes and flake production (Dibble 1985:240).

In terms of specific issues affected by raw material quality and composition, much remains to be learned. We do, however, have some important research directives which will lead to an improved understanding of how lithic types like quartzite are unique. In this way, we can make use of *external criteria* from other sciences to validate the basis of debitage analysis. Specifically, we should be exploring the principles behind fracture properties of raw materials, in terms of input conditions and output morphologies (Schurrenberger and Bryan 1985:155), as this will explain the covariance of debitage attributes and the relationship of material to knapping techniques.

Archaeologists can no longer act as if the meaning of lithic assemblages is unambiguous, since, "[w]ithout recourse to independent studies into the nature and sources of variability among debitage, such a position is not tenable" (Mauldin and Amick 1989: 85). Lithics researchers are calling, more and more, for the validation of experimental results, including studies of raw material factors. In Runnels' (1985) analysis of the state of lithic analysis, for example, his principal conclusion and number one suggestion for improvement was that lithic studies be expanded and given a broader definition that would include a wide variety of resources, such as igneous and volcanic rocks like andesite and diorite, sedimentary and metamorphic rocks like sandstone, all metalliferous rocks, and more. Until recommendations such as these are followed through with, replicative experimentation runs the risk of continuing to generate unsubstantiated and misleading claims.

The Effects of Raw Material Qualities on the Processes and By-Products of Flintknapping

It has long been recognized that many flintknapping experiments can be of limited value when they use only excellent raw materials such as vitreous and fine-grained cryptocrystalline types (Cahen and Van Noten 1971:214). This still holds true today, where the tendency is to "leapfrog" the basic-level hypothesis testing regarding mechanical properties of materials; archaeologists tend to focus on upper-level, highly theoretical issues which are less tedious and more challenging to the imagination (Tringham 1978). However, the consequences of this kind of research bias are dire, since it means that theoretical foundations are often completely unsupported and easily refuted; "each stage of the investigation must be established beyond reasonable doubt before passing on to the next" (Tringham 1978:175).

Therefore, assumptions, typologies, and theories based on limited lithic replication may not be applicable outside of the experimental environment. It cannot just be assumed that the features or characteristics of chert and obsidian debitage can also be used to identify or discriminate the reduction stage or knapping technique used on other raw materials. Crossing lithic grade scale boundaries in such a way is what Callahan (1979:24) called the most abused area in lithics research. He argued that the indiscriminate use of obsidian for replicating artifacts originally made from other materials is doubly questionable, since it is chosen mainly for its ease of flaking rather than its ability to approximate an actual archaeological scenario. Materials like quartzite, which were commonly used across all spatial and temporal ranges in human prehistory, need to be more fully explored in order to determine if their unique fracture behaviours and debitage patterns can be identified and interpreted using the methods currently available in lithic analysis.

The need to grow beyond studies of cherts and obsidian is especially true when testing popular methods like the Sullivan and Rozen typology, the Modified Sullivan and Rozen Typology (MSRT), Individual Flake Attribute Analysis, and Mass Analysis, which are all commonly relied upon in archaeological studies. In fact, all of the lithic researchers involved in the development of these debitage analysis techniques admit that more testing is necessary, especially involving an expansion of the data base to include a wider range of raw material types (Rozen and Sullivan 1989b:173; Prentiss and Romanski 1989:97; Magne 1985:255, 1989:16; Ahler 1989:219). If the results of such testing support the original claims made by the investigators, then the validity of their procedures will have been verified. If not, then we must seek to either correct any small-scale problems such as systematic errors or researcher bias, or else re-build the theoretical foundations upon which these methods rest.

F/ THESIS OBJECTIVES

To illustrate the potential difficulties inherent in debitage analysis methods, the Sullivan and Rozen typology serves as a good example. This method has been criticized by many (see Ensor and Roemer 1989; Amick and Mauldin 1989), a fact which led to an experimental re-evaluation by Prentiss and Romanski (1989), Prentiss (1993), and Ingbar et al (1989) . In essence, their findings reported that the typology was reliable, that is, it would consistently give the same results each time it was used on an assemblage (Nance and Ball 1986: 461), but it was not valid. An instrument or technique is considered valid only if it measures what it is intended to measure; in other words, validity provides an assessment of the usefulness of an instrument for actually measuring meaningful variability (Prentiss 1993:111).

Construct validity refers to whether or not an instrument's results reflect a given theoretical construct (Carmines and Zeller 1979:23: Nance 1987:287). That is, does a technique or approach really measure and explain the behaviour it is supposed to? In my research, I wanted to determine whether or not the methods designed to classify and interpret obsidian or chert debitage would be valid when used on materials, like quartzite, which fracture differently and therefore may produce different debitage assemblages and flake-class distributions. In order to assess their validity, I needed to study the extent of quartzite's own, separate diagnostic characteristics, and subsequently determine whether or not these different traits alter the validity claims for any debitage analysis techniques.

In general, I wish to discover whether or not lithic raw material type is a *controlling* factor determining debitage variability, rather than merely being something whose effects are relatively unimportant. Can quartzite debitage be "technique-specific", or technologically diagnostic for various forms of lithic reduction? Also, if these kinds of debitage patterns do exist, are they more, or *less*, pronounced with coarse-grained lithic materials like quartzite? Most importantly, I will try to determine whether or not the methods and procedures used in debitage analysis tend to be valid, both internally and externally. That is, are they valid in relation to the particular things they are measuring, and are they valid when applied *outside* the original experimental situation? Specifically, there are four main questions I want to explore and evaluate:

1. Is the identification of reduction stages independent of raw material type, as Magne (1985:127) claims? Using quartzite, I wish to find out whether or not this assumption is true, and whether or not Individual Flake Attribute Analysis is valid for lithic types like Morrison quartzite. I will therefore attempt to correlate all types of debitage to a specific reduction stage or strategy.

2. Can technique-specific debitage classes, such as biface thinning flakes or bipolar flakes, be used to identify the method of reduction that created an assemblage of flaking debris, as Magne (1985: 126-127) has argued, and are they able to accurately characterize quartzite assemblages?

3. Does the modified Sullivan and Rozen typology (see Prentiss 1993) work for Morrison quartzite debitage? That is, does it clearly delineate which reduction strategy (either core or biface reduction) was used, or do the flake categories show no patterning in this regard?

4. Are the flake-size distributions and patterns used in Mass Analysis independent of raw material type? Or, will this approach fail to demonstrate markedly different flake aggregates for the various quartzite reduction technologies and stages employed in the experiments?

To answer these questions, I will use the following general procedure in my research and replicative experiments on quartzite:

1. Explore *external* criteria such as petrology and fracture mechanics, which may indicate whether there are distinct patterns of flake breakage for different types of raw material.

2. Review the sources *within* archaeology for further insights on the validity of claims made by debitage analysts. For example, information from replicative experiments carried out in the past (along with the results from my own experiments) will indicate whether, in practice, there are significant knapping differences between various raw materials and whether these differences are patterned. That is, does quartzite knap differently and produce its own unique patterns of debitage when used in lithic reduction experiments?

3. Once these areas have been investigated, I will examine the validity of the three most common methods used in debitage analysis (i.e., Magne's, Ahler's, and the Sullivan and Rozen/MSRT approach)? If the quartzite debitage patterns show a high degree of correlation with the stage of manufacture or the type of reduction strategy predicted by these methods, then they can be considered valid procedures for interpreting debitage assemblages.

4. If the quartzite debitage patterns are not what these methods predicted, then we can assume that they are not valid approaches, and they do not measure the phenomena they

claim to. Probable causes for such a failure will be explored, in terms of specific faults of the individual methods as well as the underlying theoretical constructs that debitage analysis in general is based on.

Since one of the continuing concerns of lithic research has been to explain the factors affecting the covariance of debitage attributes, we need to increase our efforts at developing "signatures" for a broad range of potential variability in lithic debitage (Tomka 1989:157), including the poorly-understood variability caused by raw material quality. My experimental evaluation of some of the popular methods of debitage analysis will go a long way toward accomplishing this goal and toward determining if it is possible to develop a quartzite "signature". In this respect, I am continuing the quest for clear and valid relationships between debitage assemblages and the factors responsible for their characteristics. Furthermore, I am expanding the experimental realm in an effort to strengthen the unsteady foundations and frameworks currently used for interpreting prehistoric technology.

CHAPTER THREE:
THE CRITERIA FOR EVALUATING VALIDITY

Introduction

In order for lithic analysts to evaluate the methods and procedures used for studying and interpreting debitage, they need to look to external and internal sources of information to provide a more subjective test of their own assumptions. While external criteria of validation can often singlehandedly either support or negate the rationale behind the interpretive instruments or procedures, researchers usually use other sources as well, such as "internal" information derived from archaeological experiments. For lithic specialists, this type of comparison is necessary in order for their claims and applications to be credible.

A/ EXTERNAL VALIDATION CRITERIA:
PETROLOGY AND PHYSICAL PROPERTIES OF RAW MATERIALS

While not in the forefront of current lithic analysis, many researchers have stressed that, in any study of technology, we need to address the issue of raw material *properties* (Goodman 1944:416), rather than just describing area-bound lithic types, often defined on "specious criteria" such as colour (Schiffer 1979:17). We need to train more people as "anthropological lithicologists" (Knudson 1978:44), who study any and all aspects of items made of stone, and raise awareness about the importance of this aspect of lithic analysis to prehistoric people. While use-wear analysts (eg. Greiser and Sheets 1979) have contributed a great deal to our understanding of this area, the qualities of raw material such as texture, structure, grain size and composition should also be considered, together with the principles of fracture in brittle solids, *whenever* we study lithic technology. In this manner, archaeology will be in the process of "internalizing" many of these external sources of validation criteria (Nance 1987:286).

Unfortunately, most archaeologists lack the proper training to examine the physical characteristics of stone artifacts and materials, and therefore must fall back on the work already carried out by petrologists and materials scientists. The research in these "sister disciplines" focuses on the composition and fracture behaviour of rocks, often relying on abstract mechanical testing, or thin section microscopy to examine the texture, optical properties, and mineralogy of lithic materials (Clough and Woolley 1985:90; Kempe and Templeman 1983:30). Many valuable contributions to stone tool analysis have been made by workers in these areas of specialization, and archaeologists should at least make more of an attempt to familiarize themselves with these fields so that they can be used to more fully understand the validity and limitations of debitage analysis, in particular.

Basic Concepts

The vague usage of common terms like "flint", "chert", or "chalcedony" has created a vast amount of needless confusion in the archaeological literature. This is also true of the term "quartzite", which is often carelessly used in reference to many different rock types, from silicified sandstone to siltstone to the mineral quartz. It is not unusual to find confusing and misleading references such as "Morrison siltstone/quartzite" (Ingbar et al 1989:120) or "silicified sediment and solid quartzite" (Ahler 1989: 207) ; I found it extremely frustrating when trying to compare my experimental results with those of other researchers who employed such indeterminate nomenclature. If archaeologists, especially lithic specialists, were properly informed and received more thorough training in raw material identification, comparisons between collections and researchers would be greatly facilitated. The identification of lithic materials "should be the first step in any analysis or replication experiment; ...this may seem trivial, but the location of lithic sources, and the application of experimental results to study materials depends on this vital information" (Ebright 1987: 29).

Even a rudimentary understanding of some basic concepts would be beneficial to anyone studying lithic materials. Descriptions of texture, for example, are the most common ones found in archaeological reports. *Texture* refers to the size, shape, and arrangement of the component elements, and is described by expressions such as "coarse-grained" or "angular" (Pettijohn 1957:13). *Structure*, on the other hand, deals with features like bedding or stratification which are usually studied in the outcrop or geologic formation (Pettijohn 1957: 158). Unfortunately, structure is almost always ignored by archaeologists. This aspect of petrology is especially important to debitage analysis where the nature of the original raw material form can greatly affect flake characteristics such as the amount of cortex cover and the way that a material fractures. Morrison quartzite, for example, may have little or no cortex cover (as it is less frequently found in nodule or cobble forms which possess a hard, outer rind), and it often fractures along the bedding planes present in the material. Prehistoric bifaces are even known to have been knapped so as to follow the flowbands in some cherts, for aesthetic reasons (Crabtree 1967). Finally, a third important concept is *maturity*, which has to do with the degree to which a sediment has evolved from the parent materials it is derived from (Pettijohn 1957:101-102); in other words, this has to do with the degree of consolidation and bonding of the sediments. In the case of quartzite, the more mature it is, the better control it offers to the stone knapper.

Ebright (1979, 1987) is the only lithic analyst I have encountered who acknowledges the necessity of such geological training, as well as the need for standardization of terminology. Unfortunately, most archaeologists gloss over this altogether or else simply rely on vague and subjective generalizations about their lithic materials; this neglect, however, can lead to dangerous misinterpretations of data. As mentioned, the sedimentary and metamorphic varieties of quartzite may fracture and behave differently during lithic reduction, so it is particularly important to properly identify this material not only for the sake of accuracy but also because these differences will affect the choice of material, tool design and function, manufacturing techniques, and debitage production.

What is "Quartzite" ?

Unfortunately, the term "quartzite" really tells us very little, since its gradations span a wide range of composition qualities and physical characteristics. The traditional textbook definition is that it is a metamorphosed sandstone which fractures through the individual grains rather than around them (Toll 1978:47). However, this definition is both inaccurate and misleading, since it was extensively revised in 1948 to include both sedimentary and metamorphic varieties, which differ in their post-depositional histories (Preston 1988:200).

The metamorphic varieties are normally referred to as "metaquartzites" and, while they are the most widely known types, they are relatively rare (Ebright 1987: 169). The original quartz grains, which are exposed to increased temperatures and pressure, undergo recrystallization in conjunction with the silica that cemented them together (Pearl 1956:91; Preston 1988:200), resulting in a complete reconstruction of the original grain shape and size. Metaquartzites are true metamorphosed sandstone which fracture across both grains and matrix alike (Pearl 1956:91).

Orthoquartzites (the *sedimentary* variety) constitute a group that is sometimes referred to as "sandstone". They are characterized by their high quartz content (at least 90%) and silica or carbonate cement, and are well sorted and well rounded. "A quartzite is only regarded as a true orthoquartzite if ... the rock fractures through the grains rather than around their boundaries" (Preston 1988:200). They range from the immature forms (commonly called "silicified sandstone" [Ebright 1987:30]) with no alterations of the original grains, to the more mature ones characterized by changes in the original crystal lattice. Of the mature variety, Carozzi's "Type I" orthoquartzites are marked by increased grain size caused by overgrowth of the quartz grains, whereas, in the "Type II" and pressolved orthoquartzites, the grains have grown together such that the cement is no longer important as a bonding agent (Ebright 1987:30-31).

Problems in Quartzite Characterization

While each of these varieties of quartzite can be defined in handbooks and recognized by petrological experts, there is still a great deal of confusion, in practice, in the specific identification of such geological samples. For example, many of the rock types from the Morrison Formation are simply referred to as "quartz sandstone" (Wyoming State Geologist [Laramie Office], pers. comm. 1994), even though they may range from the most coarse-grained, friable varieties at the immature end of the orthoquarzite spectrum, to the finer-grained, mature materials like the type commonly called "Morrison quartzite".

This material, which was used in my experiments, is composed of almost pure quartz sand grains which are quite fine and well sorted. Based on Ebright's (1987: 30-34) summary of the identification criteria, it is a mature orthoquartzite, probably belonging to the Type II variety. However, as many researchers have also found, a transitional or less characteristic form could be easily confused (see Bridgland 1988:189; Duncan 1969:67). In fact, since most quartzite varieties cannot be determined from macroscopic examination of artifacts or hand specimens (Frison 1974: 192; Ebright 1987: 30), it would be necessary for a petrologist to analyse thin sections of this material in order to be completely certain about which specific type of the mature orthoquartzites it belongs to.

In general, the identification problem is a particularly difficult one for archaeologists to deal with, since they usually lack the proper geological training and background to fully evaluate lithic materials. Consequently, quartzite is most often lumped in with all the other coarser-grained lithic types. Artifacts are sorted into classifications like "non-obsidian" (see Boksenbaum 1980) or "non-flint", which can also include such varied rock types as slate, rhyolite, or "devitrified lava of andesitic composition" (MacRae 1988:133); therefore, such categories are essentially meaningless. Francis (1983) even lumps quartzites and cherts together under one category ("Morrison Materials"). As a result of such oversights, two negative effects are reinforced which only serve to leave lithic assemblages in a hazy analytical state. To begin with, the great amount of variation in quartzite is ignored. Furthermore, the *functional implications* of such variability, both within the class "quartzite" and in contrast to other classes, is also lost (Toll 1978:47).

Not only do researchers ignore the variability inherent in lithic type definitions, but they also assume that the mechanical properties of cherts and obsidians can safely be applied to studies of quartzite (see Luedtke 1992:91). Greiser and Sheets (1979:295), however, found that the differences in wear patterns between quartzite and obsidian or chert were significant enough that "direct extrapolation of data from one type of lithic material to another is not justified". Within the category of "quartzite", itself, there are enough differences in texture and composition to warrant individual analysis of material from each source. However, even a single formation can contain variations in grain size, texture, cohesion, colour, and mineralogical composition (Duncan 1969:39).

In Wyoming, Frison (1974; Frison and Stanford 1982:174) found this to be true of the Spanish Diggings quartzite quarries. He observed that, with the amount of variance in the lithic properties, it was extremely difficult, if not impossible, to trace the archaeological site materials to their actual sources; consequently, any criteria or characteristics used to separate the materials often lost their diagnostic value. Francis (1983) also found that all variations of Morrison quartzite could occur in the same outcrop or exposure, a fact which my own analysis of this material can support. This type of quartzite, which was used in my replicative experiments, often varied considerably from one block to the next, even though it all came from one small sector of a single exposure.

To further complicate matters, a wide range of physical properties can be encountered even within a single nodule or cobble of quartzite, exhibiting, for example, a whole continuum of granularity (Behm and Faulkner 1974:275; Toll 1978; Callahan 1979; Kamminga 1979:297;Whitehead 1988: 118; MacRae 1988: 133). Again, this is true of Morrison quartzite which can be anywhere from extremely fine-grained, to the opposite end of the spectrum (Francis 1983:148).

These problems are particularly relevant to archaeologists because they often have to rely on the results of mechanical tests to describe the petrological qualities of their samples. While some of this research is carried out by individual archaeologists, most will use the values for properties such as hardness, tensile strength, or elasticity already derived by geologists and engineers. The most prominent problem associated with this approach deals with the application of the engineering information to archaeological interpretation. Carmichael's (1989) *Practical Handbook of the Physical Properties of Rocks and Minerals*, for example, is a common reference book listing values for the properties of various lithic materials. However, since only a limited number of these rock types have been studied, one must rely on ballpark figures and values for the many classes of stone. "Quartzite" is used as an umbrella term, referring to this material as a metamorphic variety in one place, and then as an orthoquartzite in another. In addition, single samples taken from vastly different areas of the world, with vastly different geological origins and depositional histories, are often used to represent *all* quartzites. Since all properties are expressed within ranges and as averages, no consideration is given to the tremendous amount of variability that can be present not only between the varieties of quartzite (both metamorphic and sedimentary), but also within one outcrop and between sources from different parts of the world. Therefore, individual samples will no doubt deviate considerably from the average value of any property (Duncan 1969:68).

Since engineering researchers are only interested in general values, of how much use are their handbooks to archaeologists? Their tests and compilations of rock properties are useful to them for many reasons, such as locating and extracting mineral resources, or designing caverns for housing nuclear waste or underground power plants, and they also are often interested in entire rock systems or formations, rather than in small portions of the material with select properties (Touloukian et al 1981: xix-3). Nevertheless, *archaeologists* have traditionally ignored these discrepancies because the rock mechanics handbooks contain the only comprehensive body of research on this subject that is currently available to them. While some of the specialists in fracture mechanics have tried applying these techniques to archaeological material, most of the tests require specialized preparation of specimens, large quantities of the material, and a lot of time and money to set them up and run them (Greiser and Sheets 1979:294). Furthermore, the results often conflict from one analyst to the next, and from one testing method to the next. If a lithics researcher is interested in finding out a material's hardness, for example, she or he could use the Shore Scleroscope to evaluate this property, however there are at least ten different kinds of hardness tests to choose from (Hayden 1979:298), most of which produce incomparable results expressed in different notational systems; "No wonder all but the most intrepid archaeologists are inclined to flee from [this] approach!" (Luedtke 1992:81).

One solution that has been suggested (Hayden and Kamminga 1979:8), but not yet followed up on, is that we should develop our own tests, especially suited to answering archaeological questions. In this way, we may be able to bypass the redundancy and inapplicability of most current engineering techniques.

B/ INTERNAL VALIDATION CRITERIA: THE REPLICATIVE APPROACH TO STUDYING QUARTZITE

The most important source of internal information for the validation of debitage analysis is through replicative experiments which can be carried out by researchers themselves or by employing the data and conclusions arrived at by others. Regardless of which way the information is gained, internal criteria are vital to the evaluation of methods and procedures in lithic analysis since they provide a closer approximation to the archaeological applications of such methods than the external , "hard science" sources do. Debitage analysts *could*, but seldom do, rely heavily on these internal criteria to validate their approaches.

As stressed earlier, most approaches to studying lithic material characteristics and behaviour have involved the replication of artifact types and techniques using either obsidian or chert. A limited amount of experimentation has been carried out with other materials, however, such as cement blocks (Baker 1981 in Magne 1985:125), ice (Crabtree 1967), or porcelain (the infamous toilet-bowl technique often used by beginners). However, the most common materials used in replicative research, after obsidian and chert, are igneous materials such as "basalt", or vitreous trachydacite, which has been used to explore such topics as billet flaking , debitage patterning produced by different reduction modes, and reliability and validity testing (Prentiss 1993; Magne 1985; Toth 1985).

From what little experimentation that has been carried out, we only know about a few aspects of quartzite's behaviour under knapping conditions. For example, we know that it is a very stable rock type, both chemically and physically, and that it does have the predictability of conchoidal fracture. While some argue that the flaking quality of quartzite is comparatively poor (see Cummins 1983; Preston 1988:201), others claim that the finer-grained varieties show remarkably good control, similar to that of flint. Moloney et al (1988), for example, found that the greatest deterrent to successful knapping was not the hardness of the quartzite (as Behm and Faulkner [1974] claimed), but rather the internal flaws that could appear and suddenly ruin their efforts. If nothing else, quartzite is definitely agreed to be a highly variable raw material (Moloney et al 1988: Cummins 1983:199), depending primarily upon its source, morphology, composition, and texture.

The English "Bunter pebble quartzite" is widely available in the form of rounded cobbles. This variety has been used to replicate handaxes in one British project (see MacRae and Moloney 1988). Many observations were provided regarding the reduction process of these cobbles, but little attention was given to the physical factors responsible for such behaviour patterns.

Several researchers in the United States have attempted to address the underlying geological characteristics that affect the knapping of quartzite, and they approach this relationship by noting the differences between the many varieties of this material type. Crabtree (1967) commented long ago that the differences between orthoquartzites and metaquartzites become apparent when you knap with them. He found that the sedimentary types (what he called silicified sandstone) allowed more control of flaking than the metamorphosed ones, due to the more angular nature of the constituent grains. Expanding upon this idea further, Preston (1988:201) explained that mature orthoquartzites composed entirely of crystalline quartz and lacking any matrix will fracture *randomly* through both the grains and cement. Unlike flint (which consists of uniform microcrystalline quartz) and obsidian, which both produce smooth conchoidal fractures, orthoquartzites tend to have uneven fracture surfaces caused by the different orientations and sizes of the grains. This arrangement of the mineral grains, however, lends a greater average compressive strength to orthoquartzites than to that of metaquartzites (Duncan 1969:148).

In terms of general knapping behaviour, the unique mechanical properties of the various quartzite varieties will cause them to flake quite differently. According to Callahan's Lithic Grade Scale for workability (1979:16), where obsidian's value of 1.0 means that it is the easiest of the common lithic materials to knap, quartzites range from 3.5 for the fine-grained varieties to 5.0 for the coarsest ones. I would place Morrison quartzite somewhere between 4.0 and 4.5, since it meets the following criteria: it is "tough" rather than elastic like cherts and obsidian, it cannot be pressure flaked with any degree of success, and it responds better to antler billets more than to wooden ones.

Ebright (1987) also discovered in her quartzite experiments that problems in thinning often arose, and that, while some types were as easy to flake as chert or flint, others were quite tough to work. It is known that prehistoric peoples often thermally altered some lithic materials in order to improve their workability, but so far almost all experiments on this topic have used cherts; consequently, these results cannot be used to predict how quartzite would behave in similar circumstances, since the two kinds of rock are so different both chemically and petrologically. Since Toll (1978), Ebright (1987), and Behm and Faulkner (1974) are the only researchers to have published substantial reports on controlled heat treatment of aboriginally used quartzites, the data base is somewhat limited. This is especially problematic considering the great variability in the behaviour and physical properties from one type to the next.

Despite the amount of experimental research on heat treating, little is known for certain about what exactly causes the desired changes in the rock. It is thought by some that the thermal alteration of cherts is related to an artificially-induced fluxing of the matrix which acts to "homogenize" it by removing the voids between grains; as a result, the chert will fracture through the microcrystals (Mandeville 1973). According to another theory, water escapes during the heating process, causing micro-cracks to form throughout the material; this serves to reduce the tensile strength, making it easier to knap (Flenniken and Garrison 1975; Ebright 1987:33).

With quartzites, only silicified sandstone and the Type I mature orthoquartzites will respond to thermal alteration. These *sedimentary* types are thought to undergo a fluxing similar to that which happens to chert. The metaquartzites, and pressolved and Type II orthoquartzites, on the other hand, are not subject to fluxing due to their nearly pure quartz content and lack of cement. In these types, either metamorphic recrystallization, pressure solution, or secondary overgrowth have already caused the grains to interlock in the same way that fluxing would, so their workability is not significantly improved by heat treatment. Thus, the response of quartzites is directly related to their maturity and to the quantity of cement present (Ebright 1987: 36).

In practice, then, there will be a range of responses in quartzite, depending on its geological history. Toll (1978) reported that the silica cemented, sedimentary variety responded well to heat treatment while another, more mature material fractured and functioned the same as when unheated. In transitional types, researchers have found that the quartzite was *somehow* changed, although they were unable to report any specific improvement (Behm and Faulkner 1973; Flenniken, pers. comm. in Ebright 1987:33). This no doubt happens because there is always a great deal of variability and overlap both within and between quartzite sources, in relation to the formation processes that created the material. Each type will therefore have unique quantities of cement, trace minerals, and pressure solution that determine how it will behave during artificial thermal alteration.

My preliminary experiments with thermal treatment of Morrison quartzite provided further support that this material is a Type II or transitional variety of mature orthoquartzite. After eight hours in a sandbath at 800° C, some colour transformation was evident (a slight change to a pink or orange hue was noted on some of the samples), probably due to increasing oxidation of iron impurities in the rock. However, the material did not demonstrate a reduction in the effort needed to remove a flake nor any improvement in general flaking control using hard hammers, billets, and pressure flakers.

Only when more experiments are conducted, using different types of quartzite, will we be able to draw any definite conclusions on why it behaves the way it does, both in heat treating studies and in the replication of a range of artifact forms.

C/ DESIGN THEORY

The main criticism that many researchers have about functionalist interpretations being too limited or distanced from the "real world" is true in many cases, such as laboratory studies which become an end in themselves or which are inapplicable to the anthropological side of archaeology. What we need to do is to incorporate aspects of both functional and behavioural perspectives in our interpretations.

One field of lithic analysis that strives to be more holistic in its approach, by dealing directly with the range of decision-making processes involved in the selection and use of lithic materials, is known as Design Theory (see Hayden and Kamminga 1979:6). The main goal in this type of research is to identify the factors affecting lithic variability and assemblage composition (Pokotylo and Hanks 1989:49). According to Hayden et al (1993), this is accomplished by examining a variety of "constraints" involved in solving problems by technological means. These are organized into five main areas of consideration, including technological, socioeconomic, and prestige constraints, as well as raw material and task performance concerns. My research into external validation criteria will focus on the constraints affecting raw material selection.

Factors Influencing Raw Material Selection

1. Availability

Until more experimental work has been carried out, we must rely on other sources of information for our insights on many issues related to lithic analysis. One such source is the archaeological record itself, which can be used to generate models for interpreting prehistoric stone use.

Based on fieldwork all over North America, we know that the main sources of prehistorically-exploited quartzite are in the Rocky Mountain region (eg. Cypress Hills and Stony Plain, Alberta; Spanish Diggings, Wyoming; Bridger Basin, Wyoming and Utah), along the Eastern seaboard (eg. South Bay, Maine; Robesonia, Pennsylvania; Piney Branch, D.C.), and around the Great Lakes (eg. Silver Mound, Wisconsin; Sheguiandah on Manitoulin Island, Ontario) (Ebright 1987; Callahan 1979). The wide distribution of quartzite made it a viable option for early stoneworkers, particularly when faced with a lack of other useful sources like chert and obsidian. Being a granular rock, quartzite does not produce the razor-sharp edges that these kinds of material do, however, it tends to be tougher than most. Morrison quartzite itself can be quite fine-grained and, although it can lack the structural strength of cherts, it was still favoured for the production of a range of projectile points and tools from Folsom times up to the historic period (Frison and Bradley 1980:12).

Little discussion is usually devoted to the *meaning* behind the raw material preferences found in the archaeological record. Most of the time, researchers will mention which raw material types are present in the assemblages, but without any hypotheses about why certain ones were employed (see Frison and Bradley 1980; Wynn and Tierson 1990). In many cases, the differential use of raw materials merely reflects a difference in the lithic resources available in each region, rather than some form of regional cultural preference (Simpson 1990:126). Unfortunately, many people then assume that alternative lithic sources were used only when the "good stuff" ran out, rather than considering the valuable contributions that each type made to various aspects of technology. Once again, as in replicative experiments, we see the bias toward chert and obsidian taking its place in archaeological research. This narrow perspective merely causes us to ignore the possible changes in lithic technology that were caused by differences in the flaking properties of available raw materials (Hiscock 1986:42).

2. Petrological Aspects of Choosing Useful Stone

Some would argue that physical differences between raw materials have only a small influence on how stone tools were made and used (eg. Luedtke 1992:80; Wynn and Tierson 1990:81), however, I believe that this attitude grossly underestimates the importance of lithic quality and mechanical properties, and is at odds with the findings of many lithic analysts.

Use-wear analysts have recognized this for some time now, and they have contributed a great deal of comparative research involving the testing of different raw material types. The degree of edge crushing and microflaking, for example, is now known to vary from one material to another as a direct result of different physical properties and petrological compositions between lithic resources. These differences are quantifiable on both actual artifacts and experimental specimens. In quartzite, they become manifest in the shallower scarring and the inconsistent length and direction of flakes, as compared to obsidian (Greiser and Sheets 1979:290-292; Toll 1978).

Crabtree (1967) and others have worked to make the study of lithic resources a standard part of archaeological analyses by explicitly examining certain desirable physical properties and features of stone materials. A brief summary of these salient traits is provided here.

Texture is the most important key to the workability of lithic materials since it can indicate: how much force is necessary to remove a flake; whether the striking platforms will collapse; whether the material can be flaked by pressure or percussion, and so on (Crabtree 1967:23). Texture, which is determined by the fineness or coarseness of the microcrystalline structure (that is, the size and shape of the constituent grains), as well as by the consolidation and bonding of the material, is the first concern of a knapper, because the coarser the texture, the harder it is to remove regular and uniform flakes (Crabtree 1967:8).

In order for a stone knapper to be able to control the thickness, width, and length of the flakes removed, the material must demonstrate a certain degree of homogeneity (that is, it must be of uniform composition, free of flaws, inclusions and cracks which cause the fractures to be deflected as they move through the material), otherwise it will break unpredictably and cause hinge fractures (Brown et al 1982:5; Luedtke 1992:86). Inhomogeneities can sometimes be detected by judging the sound of a nodule or core; if it gives off a dull sound, this is because the velocity of the sound waves is affected by the presence of cracks, fissures and planes of weakness (Crabtree 1967:9). Australian aboriginals use such a technique to judge the quality of quartzite (Binford and O'Connell 1984).

Isotropy is not the same as homogeneity, although it is often confused with it. According to Luedtke (1992:86), the mechanical properties of a material must be the same in all directions (that is, it must have the same strength no matter how it is oriented) for it to be considered isotropic. This property is missing in rocks like shale which have definite fracture planes, but it also affects raw materials like chert which can be anisotropic due to fossil inclusions. Some of the quartzite used in my experiments was discarded because its isotropic value was ruined by large internal fissures and flaws, laminations or impurities that had accumulated along the deposition planes, and frost fracturing caused by being exposed on the ground surface (see Luedtke 1992:86).

Strength (also called Toughness or Tenacity), which measures the amount of resistance to the necessary force required for detaching a flake, is another of the properties affecting aboriginal stone working such as percussion flaking. Often defined as the quality of flexibility without brittleness, or yielding to force without breaking (Crabtree 1967:24), strength is exhibited by materials such as flint, chert, jasper and some fine-grained quartzites. However, it is not a property of brittle obsidians at the low end of Callahan's lithic grade scale, nor by coarser-grained quartzites, rhyolites, slates or common basalt at the opposite end. There is no single measure of strength because materials respond in a number of ways to different kinds of stress such as compressive or tensile (Luedtke 1992:88). However, in general, the factors that determine strength are mineralogy, grain size, and the extent to which the grains are interlocked (Luedtke 1992:88). Consequently, this explains the wide range of different values for strength or toughness between the various types of quartzite, such as those reported by Goodman (1944). Coarser-grained materials tend to have larger flaws along grain boundaries and more cracks and pores which weaken a rock considerably and also make it very prone to fracture (Luedtke 1992:88).

Lower strength may be the reason that the "Shirley Basin quartzite" biface produced far more debitage than the "Morrison siltstone/quartzite" one did in the experiments carried out by Ingbar et al (1989); this may also explain why it "died" of an endshock break. However, lack of elasticity, which is the ability of a lithic material to resist unwanted fractures such as endshock or step and hinge fractures (Luedtke 1992:90), may also have been at fault. This property, or complex of properties, is correlated with hardness and is also affected by the presence of cracks and the size of the constituent grains. Fine-grained rocks tend to be more elastic, or more able to return to their original form when the force is released (Crabtree 1967:24), although the measures of resiliency equated with this property, such as Poisson's Ratio, vary from test to test. In general, elastic materials allow a person to control and guide a flake over a curved surface. According to Crabtree (1967:24), "Heated cryptocrystalline minerals and volcanic glasses have this flexibility to a greater extent than the coarser materials [and a good] flintknapper can control the flexing to an amazing degree".

Finally, hardness, or the ability to resist abrasion, scratching, or penetration by an indenter (Luedtke 1992:91), is another of the complex properties that affect the behaviour of a rock type during lithic reduction. Smaller-grained rock types have greater hardness than the coarse varieties, although specific values for hardness will change depending on which of the many tests are used. On the Moh's scale, which is only useful for gross differentiation, quartzite is not given an exact value (Toll 1978:49), probably because the range is too great, although personal testing of Morrison quartzite places it in the 6.5 to 7.0 range. Like strength and elasticity, it is really a composite property affected by various geological factors.

As seen here, there are a number of important areas of agreement between the mechanical properties such as hardness or elasticity and a raw material's petrological traits such as texture and grain size; in fact, all of these concerns reflect what experimental knappers and modern stoneworking groups consider the best knapping qualities of a lithic resource (Luedtke 1992:114).

3. Aesthetic and Symbolic Factors in Raw Material Selection

In the past, raw material choice may not always have been related to availability of particular types or to mechanical advantages. Some archaeologists argue that our perspectives on this issue are too narrow or biased by our scientific attitudes and approaches. Instead, many have suggested that we try to go from the practical aspects of raw material choice to understanding the value system that was also influencing lithic selection; "is there ever an artifact produced that doesn't have symbolic loading?" (Kleindienst 1979:299).

Choice of materials may well have been limited by availability or utilitarian considerations, but factors such as tradition or ideology could also have been an important part of the decision (Goodman 1944; Gould 1978; Pokotylo and Hanks 1989:54-55). Prestige considerations, as well, may have been crucial in the selection of raw materials, however Hayden et al (1993:42) argue that this would only happen in rare instances.

Aesthetic reasons may also have been an important determining factor. For example, some gold coloured jaspers turn bright red when heated, and Snow (1980:132 in Luedtke 1992:103) has argued that such a magical change to the colour of blood might have been ideologically significant to prehistoric hunters of the northeast. Luedtke (1976) also suggested strong ritual associations between certain strikingly-coloured cherts in Michigan during the Late Woodland. In addition, Gould (1978:285-287) has noted similar ideological connections between some brightly coloured adzes and circumcision knives among the Australian aboriginals. These lithic materials carried a great deal of prestige and mystical power since they were obtained from specific geological sites affiliated with sacred ancestral Dreamtime characters. Some quartzites, in particular, are said to be spiritually charged with power that will make spear points more effective in hunting and warfare, since their brightness and iridescence is directly connected to the Ancestral Beings (Taçon 1991). The aboriginals in Arnhem Land chose to exploit certain quartzite outcrops because they were believed to be the petrified remains of these characters. Some researchers argue that we need to use such ideologically-oriented approaches because strictly functionalist theories tell us relatively little about the significance and meaning of lithics (Taçon 1991; Gero 1991).

Sometimes there is simply a lack of understanding of *why* certain rock types are chosen, or perhaps it is difficult to identify which particular attribute makes a material useful or desirable. For example, some Australian aboriginals could not explain the functional aspects of why they preferred one type over another when asked to specify the reasons for choosing certain lithic materials (Gould 1978). This, in turn, may be linked to the fact that the physical characteristics of stone overlap to such a degree. Even when petrologists use terms like toughness or hardness, they are really describing a *complex* of traits; therefore, it is difficult to pinpoint which aspect (if any) is responsible for the behaviour of a material (Hayden 1979:298). Texture and porosity depend on the shape and sorting of grains (Pettijohn 1957:13), while hardness is affected by strength, elasticity, and cleavage. As a result, many mechanical properties are strongly correlated across rock types such that strong materials, for example, are usually also hard and stiff (Luedtke 1992:91-94). Given the complexity of both the properties and the task of sorting them out, more research into the relationship between toolstone selection and functional considerations is required, particularly beyond the realm of usewear studies.

While it *is* important to remember that lithic items were part of a larger cultural and belief system, we must not condemn or negate the practical, utilitarian determinants of raw material selection; rather, we should seek to endow this aspect of ancient technology with a broader frame of reference. Most archaeological tools were practical technologies and therefore we can expect practical considerations to be most important in initial modeling.

4. Functional Aspects of Quartzite Tool Design

Archaeologists have long noted the differences between lithic types in their collections. However, these proportional differences are seldom considered in light of the physical and mechanical properties of the stone itself. From the analysis of lithic assemblages around the world we know that, prehistorically, certain types of stone were preferred for certain tasks, primarily because of how these physical features caused them to behave during knapping and use. Most people assume that a prehistoric stoneworker's knowledge of the physical properties of lithic resources must have influenced the deliberate selection of certain rock types for certain purposes.

Quartzite, in fact, was a " highly functional lithic material, if not a preferable one for some tasks" (Toll 1978:64), as evidenced in the differential use of it for various tool types; in this way, the relationship between the lithic material, tool design, and function is clearly demonstrated (Hayden 1979:298-299). Archaeological assemblages have shown us that quartzite was the preferred lithic material for certain kinds of flake tools such as notches, gravers, denticulates, and scrapers (Greiser and Sheets 1979; Crabtree 1967, 1975; Toll 1978; Frison and Bradley 1980; Strauss 1980:70; Francis 1983; Beck and Jones 1990) as well as for sickle blades in the Near East (Cummins 1983:207). In addition, quartzite provided a coarse, granular edge that was well suited for activities such as the working of bone, wood, and antler (Crabtree 1967:11). Whereas the razor-sharp edge of an obsidian flake provides an extremely effective tool for cutting soft materials, the same thin edge will not stand up very well to rigourous scraping or working of very hard substances like bone or wood, since it cannot sustain its sharpness over a long period of use. Less brittle materials such as chert and quartzite tend to retain sharp edges longer and, once they become dulled, they are easily resharpened by secondary flaking (Cummins1983:203).

Sharp, *un*modified quartzite flakes were preferred in a number of circumstances, and often were considered to be as important as retouched flakes and tools, especially for activities related to animal butchery (Toth 1985:109). A preponderance of such unretouched quartzite flake tools has been noted around the world, from Europe (Rankama 1990) to Africa (Davis 1978; Toth 1985), Australia (Gould 1978; Binford and O'Connell 1984), Britain (Inskeep 1988) and North America (Lee 1954,1955; Toll 1978;

Frison and Bradley 1980; Francis 1983; Ebright 1978, 1987), to name but a few examples. Quartzite was more often used in unmodified flake form than were other materials because of the durability and toughness of its edges, its less workable nature in general, and larger size of quartzite debitage relative to other fine-grained materials (Toll 1978:54).

The toughness of quartzite also made it useful for activities involving shock or great stress (Strauss 1980:70), such as chopping tools, knives, and spears. Some archaeologists have also observed that many of the coarser types of quartzite are not employed in the manufacture of refined implements such as thin bifaces or pressure flaked tools (Crabtree 1967; Bradley and Frison 1980; Cummins 1983; MacRae 1988). However, Lee (1955) observed polyhedral blade cores made from Munising orthoquartzite at the Sheguiandah quarry in Ontario, and Frison and Bradley (1980) have noted the use of quartzite in very finely retouched fluted points in Wyoming.

Nevertheless, the conclusions of most experimenters support the observations by archaeologists regarding the unsuitability of some quartzite for refined thinning operations (eg., Jones 1979; Ebright 1987; Moloney 1988). In general, the mechanical properties and coarser grain size have been found to restrict the secondary thinning and pressure flaking of quartzite tools, such that bifaces, for example, tend to have a rather crude appearance. As Crabtree (1975:109) warns, however, we should not indiscriminately write off such tools as being the result of inferior knapping skills; " if you consider the material, you may discover that you have a superior [stoneworker] who was forced to use inferior material".

It would appear, then, that availability, ideology, functional demands, and tool design are intricately connected factors capable of influencing raw material selection on a number of levels. Sometimes one factor will overshadow another, depending on situational concerns; consequently, a certain material may be good from a manufacturing perspective but not from a usewear one. Greiser and Sheets (1979:295) use the example of how modern knappers often insist on using obsidian because it flakes so easily, despite the fact that the obsidian flakes and tools they tested were the first ones, out of a variety of material types, to lose their working edges through rapid attrition. Once again, we see the bias toward obsidian in lithic analysis. Unfortunately, in this case, this bias can have the effect of keeping other kinds of raw materials out of the experimental realm, even though they may have contributed more to actual prehistoric technology and subsistence practices.

D/ HOW RAW MATERIAL PROPERTIES MAY AFFECT THE VALIDITY OF ARCHAEOLOGICAL ASSUMPTIONS

The bias in experimental lithic analysis against materials like quartzite can be shown to have a detrimental effect on the models we develop for interpreting the past, when these models are based only on a non-representative material like obsidian. The discrepancies between the physical properties of obsidian and those of the rest of the suite of exploited lithic resources are significant enough to possibly create invalid correlations between fracture features and reduction mode, for example, or to provide inaccurate identifications of manufacturing stages, if these models are used to interpret tools and debitage made from non-obsidian materials.

To determine why and how this is true, it is helpful to "borrow" from an outside scientific discipline so that we can examine the very fundamental principles governing the way a lithic material fractures. These underlying principles can then be used in conjunction with the other external validation criteria (the physical properties of the material) to come to an understanding of the characteristics of debitage formation. However, fracture mechanics is still poorly understood by most and is seen as unapproachable to all but a few who have received training in engineering or materials science. Consequently, we find that most archaeologists lack familiarity with both fracture principles and the conditions affecting flakability, and yet they will still incorporate these topics in their interpretations of lithic assemblages (see Hiscock 1986). The application of partially-understood fracture mechanics principles can result in misleading explanations of lithic manufacturing processes and by-products.

Although certain aspects of fracture mechanics do vary according to raw material type (as in crack velocity, for example, which is twice as fast in obsidian as it is in quartzite [Cotterell and Kamminga 1987:681]), they may have little or no effect on the debitage produced during reduction. On the other hand, raw material type can play an important role in creating or exhibiting fracture characteristics such as ripple marks, hackles, or Wallner Lines, which are used to differentiate the various kinds of fracture which formed a flake (Cotterell and Kamminga 1979).

Although commonly found in both archaeological and experimental debitage assemblages, it is possible that these features will not even be manifested on some of the more coarse-grained materials (Hayden and Kamminga 1979:7). In terms of quartzite, certain varieties are said to show other diagnostic traits such as platform lipping, bulbs of percussion, eraillure flakes, or Hertzian cones in varying degrees, which makes it more difficult to reconstruct manufacturing behaviour from the waste products (Patterson and Sollberger 1978:107; Watson 1983:123; Ebright 1987:30; Inskeep 1988: 239; MacRae 1988:136; Moloney et al 1988:39; Hayden and Hutchings 1989:241). As a result, archaeologists are sometimes unable to classify non-obsidian or non-chert debitage with any degree of certainty (Boksenbaum 1980:16) and must make judgment calls when recording flake features such as dorsal scar counts (Odell 1989:194).

Aside from texture or granularity, many other aspects of physical structure will have a very significant influence on the formation of flakes, depending upon which material is being knapped. For example, tensile stress (which is the most important kind of stress in a flaking situation) is directly related to the brittleness of a lithic material; the more brittle a stone is, the more likely it is to break when stress is applied. In other words, "the fracture strength of a material should be equal to the amount of stress needed to break its weakest atomic bonds" (Luedtke 1992:82-83). The atomic bonds, in turn, are determined by the physical or petrological structure of the material. Fracture strength is influenced further by the presence of inhomogeneities, such as tiny Griffiths cracks on the surface, which disturb the propagation phase of flake formation by deflecting the fracture as it moves through the material (Cotterell and Kamminga 1987: 698)

In this way, "the overall shape of a fragment produced by fracture is controlled by the distribution of stresses within the core or tool" (Luedtke 1992:84). This fact has important repercussions for debitage analysis, because it argues that certain aspects of fracture mechanics are different for each lithic material type, and therefore the debitage produced will be different for each type as well; that is, each different raw material will produce its own types of flakes (Newcomer and Karlin 1987:36).

This claim has been supported by the preliminary findings of Ingbar et al (1989), whose replication project employed several types of lithic materials, including the only published example yet of experimentally-evaluated Morrison quartzite. They found that the quartzite debitage distributions produced during the manufacture of a biface were unsuccessful 50% of the time at identifying the reduction technique that had been employed in the test. Nevertheless, they were somewhat vague in their explanations of what was responsible for the differences in flake breakage frequencies, giving raw material differences an equal consideration to other factors such as core form, skill levels of the knappers, and the different reduction techniques used (Ingbar et al 1989:121).

Perhaps the main reason for such uncertainty is the fact that raw material variables are hard to quantify, despite the acknowledgment that they undoubtedly affect flake morphology (Dibble 1985:237). Sullivan and Rozen (1985:769), however, firmly believe that the proportions of flakes in their debitage categories are related to the physical properties of the material which are responsible for mechanical failure, or fracturing of the stone; therefore the type of reduction is secondary to the petrological nature of the raw material. Their claims were tested by Prentiss and Romanski, who observed how inhomogeneities in chert, such as incipient fracture planes, will produce more broken flakes than a material that was flawless; this would increase the number of "Proximal Flakes" in the Sullivan and Rozen typology when applied to the debitage of an inhomogeneous material. In essence, their experimental findings confirmed that lithic "fracture properties will greatly affect flake breakage and the subsequent debitage assemblage patterning" (1989:93).

Prentiss' (1993) work is unique in its thoroughness and scientific stringency and is of particular note since little experimental evaluation of debitage analysis methods has been carried out. As we have seen, using unsubstantiated claims or typologies can produce questionable results when applied beyond obsidian or chert debitage. Magne's approach, as well, was successful for one type of material, but even he acknowledges that "the classification of debitage here is certainly in need of independent verification" (1985: 254). His concerns are echoed by Ahler (1989:219) who called for an expansion of the database to include experiments in a wider array of lithic material types, following testing that focused principally on Crescent Chert and Knife River Flint. It is possible that his size-grading approach also runs the risk of being ineffective when used on lithic types other than cherts, since the size of flakes can be influenced by the composition of each material. The far-reaching effects of a raw material's behaviour and variability on lithic analysis are only just beginning to be appreciated as we continue to consider the external validity (or "generalizability" [Moore 1991]) of debitage patterning outside of the original experimental settings.

E/ SUMMARY

This chapter has shown how differences in geological history, purity of composition, granularity, and physical properties such as elasticity and strength can affect the way that a lithic material fractures during reduction. Consequently, these factors will influence the workability of the stone, the choice of raw material, and ultimately the finished product and its associated debitage. While some experimental evaluation of the various methods used to analyse debitage has been undertaken, this has so far been at a preliminary or single-case level only. All researchers in the field of debitage analysis agree that the relationships between petrology, fracture form, and raw material are hazy enough to merit further investigation using a wide variety of lithic resources. It is my hope that my quartzite experiments will make a valuable contribution to our understanding of this subject.

Since the physical aspects of a lithic material have such an effect on the shape, size and other fracture characteristics of debitage, it makes sense to explore this relationship further because it will determine whether or not the methods and procedures used in debitage analysis are in fact valid instruments. My research project is designed to observe the behaviour of quartzite under controlled replicative testing and to determine how its physical and petrological characteristics affect the morphology of debitage. This approach will continue to explore the internal validation criteria used to describe debitage, by determining whether or not there is a unique pattern to quartzite reduction by-products. By comparing these findings to those from studies of obsidian and chert, I should be able to answer the second of my research concerns, which asks if debitage patterns are independent of raw material type. The next step, then, is to test the validity of the typologies and approaches used to study debitage. That is, I will be asking whether or not they measure the phenomena they claim to. Subsequently, I will also be able to determine the extent to which the theoretical framework of these analytical techniques, and of debitage analysis in general, is a valid interpretive tool.

CHAPTER FOUR:
DATA COLLECTION AND EXPERIMENTAL DESIGN

Introduction

This chapter describes the background to the research questions I have posed and the methodology I have designed for addressing them. Specific details of the data collection procedures and the experimental aspect are therefore provided, together with the identification of possible confounding variables that could negatively affect the results. Finally, I will discuss any expectations that I had, going in to the experiments, regarding the interpretation of quartzite flaking debris.

I have decided to assess the validity of three common types of approaches or procedures used in the analysis of lithic debitage: 1. Individual Flake Attribute Analysis, as employed by Magne (1985), 2. the Item Completeness method developed by Sullivan and Rozen (1985; Sullivan 1987) and expanded upon by Prentiss (Prentiss et al 1988; Prentiss and Romanski 1989; Prentiss 1993), and 3. Mass Analysis (after Ahler 1989a, 1989b). Based on the information obtained from internal and external sources (the insights gained from petrology, fracture mechanics, and replicative experiments discussed in Chapter Three), I suspected that the effects of differences between raw material types would have a detrimental impact on the ability to identify technological origins using debitage methods developed and tested with only vitreous or fine-grained microcrystalline material. For example, if I were to use Magne's method to analyse *quartzite* debitage, would the scarring patterns that he used to define specific reduction stages be the same on these flakes, even though it has been shown that quartzite has its own distinct morphological traits and may not even show such fracture features very clearly, if at all?

A/ WHY THESE METHODS MAY NOT BE VALID

How They Operate

1. Individual Flake Attribute Analysis

Since this type of debitage analysis has been the most commonly employed approach, I was interested to see how one of these more traditional procedures would classify flakes and subsequently interpret the data derived from the typology. As mentioned, I have chosen to focus on Magne's (1985) scar count method for interpreting debitage assemblage formation.

To begin with, this classificatory procedure relies on the concept of sequential reduction stages, which he divides into "Early", "Middle", and "Late" parts. To accomplish this goal, Magne separates the flakes into a stage classification through the quantification of dorsal scars (for "Shatter", or flakes without striking platforms) and platform scars (for "Platform-Remnant-Bearing Flakes", or PRBs). Two distinct reduction types are also identified according to the presence of pre-determined attributes thought to be technique-specific. These two types include the biface reduction flake and the bipolar flake, which essentially are specific or specialized subdivisions of the PRB and Shatter categories, respectively (Magne 1985:129).

This method developed out of an earlier pilot study (see Magne and Pokotylo 1981) that examined the debitage resulting from the manufacture of flake blanks and a single biface. Its purpose was to use multivariate statistical techniques to narrow down a long list of debitage variables to a shorter one that would eliminate redundancy in attribute measurement. The six variables they arrived at were then incorporated in Magne's doctoral research (published in 1985 as *Lithics and Livelihood: Stone Tool Technologies of Central and Southern Interior British Columbia*), and he was able to shorten the list even more, yet still allow for an accurate means of distinguishing block core reduction from biface manufacture. The taxonomic classification scheme includes the key flake types mentioned above, and is relatively simple and quick to use.

2. Item Completeness Approach

Some of the problems associated with the individual flake attribute analysis were cause for concern to a number of researchers, although there were few alternatives available until the mid-eighties. Sullivan and Rozen (1985; Sullivan 1987), for example, were not convinced that individual "technique-specific" flakes could be used to successfully infer the technological origins of a large and variable debitage collection. Rather than making such inferences at the artifact level, they argued that entire assemblages should be examined, without being linked *a priori* to specific conclusions based on the presence of debitage types such as "biface thinning flakes" or "bipolar flakes" (Sullivan and Rozen 1985:759). Their "interpretation-free" approach, based instead on the proportions of flakes in each of the four mutually-exclusive completeness classes, is said to be able to distinguish core reduction from biface manufacture. Specifically, they argue that core reduction produces high proportions of complete flakes and nonorientable "debris", whereas biface manufacture produces large amounts of flake fragments (the "medial-distal" pieces without striking platforms or intact margins) and broken, or "proximal", fragments which result from the breakage of very thin biface reduction flakes (Sullivan and Rozen 1985:769).

Since this approach was derived from archaeological rather than controlled experimental means, Prentiss and Romanski (1989) carried out their own experiments to test this approach. They found that, while it was able to differentiate core reduction from "tool", or biface, manufacturing debitage, it did not take taphonomic factors such as trampling into account, even though trampling greatly alters the percentages of flakes in each category. Prentiss (1993) continued his experimental evaluation of the typology and found a way to improve its validity. By adding size grades to the basic flake breakage classes, he was able to obtain results that were much less biased by factors such as trampling. The size categories were: "small" (less than 4 square cm), "medium" (4-16 square cm), "large" (16-64 square cm), and "extra-large" (greater than 64 square cm) (Prentiss 1993:74). He called this version the "Modified Sullivan and Rozen Typology", or MSRT, and then applied it with favourable results to an archaeological collection from the Keatley Creek site in British Columbia. I wish to further explore the apparent success of this new typology when applied to quartzite debitage. Would the flake class distributions be the same as they were with obsidian, or would the typology not be able to differentiate quartzite core reduction debitage from bifacial tool manufacturing debris?

3.Mass Analysis

By adding the size grades to the Sullivan and Rozen typology, Prentiss's modified version becomes an intermediate between individual flake analysis and the final analytical approach I will be studying, which looks at size grades of entire debitage assemblages. While other researchers (eg. Stahle and Dunn 1982; Patterson and Sollberger 1978;) have employed this type of procedure to debitage analysis, Ahler (see 1989b for a history of the approach) has remained in the forefront by consistently testing and expanding his experimental database for this technique over the last twenty years or so.

Essentially, Mass Analysis is a branch of Flake Aggregate Analysis which studies all the flakes in a debitage assemblage and their proportions within four or five size grades. Entire flake collections are passed through a series of nested screens, each lined with mesh or hardware cloth that has a different sized opening than the others. In general, the standardized sizes include one inch (G1), 1/2 inch (G2), 1/4 inch (G3), and 1/8 inch (G4) openings. A few select variables are also collected for each size grade, including total number of flakes, counts of cortical flakes, and the collective weight, although no attributes are recorded for individual flakes.

The basic theoretical premise is that various stages and techniques of reduction will produce distinctive distributions of flakes across the size grades. These patterns can then be used to interpret other debitage collections in a reliable, fast, and straightforward manner. The initial screening, weighing, and counting is followed by simple calculations of mean flake weights or frequency ratios, for example, which will provide additional technological information that is inherent in the basic data set.

Mass analysis is unique in its simplicity, and appears to be even more free of interpretive bias than the Sullivan and Rozen or MSRT methods. As such, it is attractive both in practical ways and in terms of its ability to obtain objective and reliable results. However, like the other approaches, it is still in need of experimental verification and validation.

Reliability and Validity in the Three Methods

While they may seem able to offer successful ways of deciphering the variability present in flaking debris collections, there are a number of areas where these three methods fall short. In general, before, any of them can be used with much confidence, it must be determined with further testing just how internally valid they are, as well as how applicable their results are outside of the original experimental setting. Furthermore, their theoretical basis must be validated to ensure that the percentages and distributions of flake types actually reflect meaningful patterns in the data. To accomplish these goals it is necessary to explore the topics of reliability and validity in more detail, with respect to the specific methods that I will be assessing. This is an important prerequisite to my experiments because it is vital to gain a sense of their strengths and weaknesses prior to applying them to my quartzite flaking debris. In this way I will be better prepared to understand *why* they succeed or fail to provide valid interpretations of a different raw material type.

Reliability is related to the consistency and replicability of a measuring instrument or approach, and is a standard of how precisely procedures are carried out. Unambiguous instructions, and explicit definitions of attributes or other concepts enable experimenters to achieve the same results each time they are repeated (Knudson 1983:6). Random error, however, may affect empirical measurements in a negative manner as a result of a general lack of precision in the instrument or in applying it . This kind of error, which is associated with reliability rather than validity, is used to designate any chance factors or unsystematic errors such as mistakes in coding, ambiguous instructions, analyst fatigue and so on; these random errors are expected to cancel each other out in the long run (Carmines and Zeller 1979), although they can be serious enough to reduce the value of an instrument because of a lack of consistency (Amick et al 1989:3).

In terms of the three debitage analysis methods being studied here, reliability problems can affect them to varying degrees. For example, all three could be affected by random mistakes made while carrying out the procedure, or by errors made during the recording of the data, although these are possible in virtually any analytical situation.

Of the more destructive kinds of random error, however, the approaches requiring higher numbers of decisions or subjective estimations (eg. of amounts of cortex cover) tend to have greater difficulties producing reliable results. This is not particularly encouraging for anyone using attribute or item completeness analysis, as Odell (1989) discovered in his study of reliability in debitage analysis. In fact, he found that no lithics researchers had yet been able to establish a credible level of replicability in their measurements, a consideration that is crucial in this field . For example, when an analyst must record the number of dorsal flake scars, he argued that "the difference between three and four ...does not sound like much until one realizes that the error is in order of 25%" (Odell 1989: 167). He found that reliability problems were pervasive due to the low levels of replicability encountered whenever measurements and estimations of variables are required by an analytical approach.

In terms of the three specific methods I am studying, this means that Magne's approach is most likely to be detrimentally affected by random error, since his approach primarily involves counting flake scars on both the dorsal surface and on the striking platform, as well as identifying technique-specific flakes based on a somewhat subjective assessment of flake features. Approaches like this, which base their flake types on a number of variables, "force the analyst to make a "best fit" assessment of the combined expressions of these variables in order to classify artifacts" (Rozen and Sullivan 1989a). Magne (1985: 114) himself found that there were many instances when it was too hard to count the scars on thin platforms, or when so-called biface thinning flakes were created by core reduction techniques. Consequently, his classifications were occasionally based on guesses or formalized definitions that may not have approximated the actual reduction event. This type of problem with reliability will occur because the measurement is somewhat imprecise and might not be repeated in the same way by another analyst.

The same problem could be experienced when applying the Item Completeness methods, however, since it is sometimes difficult to determine which category the flake should belong to. The Sullivan and Rozen criteria are often ambiguous and hard to apply in all situations, particularly with "atypical" or transitional flakes. Furthermore, raw material factors such as grain size may obscure the attributes used in their typology, particularly the identification of the ventral surface or the point of applied force, thus forcing the same type of "best fit" classification that they were trying to avoid in the first place.

Mass Analysis, on the other hand, has tried to eradicate the subjective side of debitage analysis by eliminating the guesswork and reducing the technique to straightforward procedures like counting and weighing. Furthermore, instead of *estimating* the amount of cortex cover on a flake (often based, in other approaches, on specific percentages or increments of cortex), this method merely involves the counting (in each size grade) of all flakes with any amount of outer rind or original outer surface. So, unless the analysts can't count or read a scale, this method allows the results to be much more reliable and repeatable. Ahler (1989b:205-207) reported that reliability testing of this method (using discriminant functions analysis) showed correct flake classification rates ranging from 79% to 99%.

Validity, on the other hand, is often determined by how much non-random, systematic error (or, "bias") takes place when the method is applied. Erroneous interpretations and conclusions can occur in debitage analysis mainly because of differences among analysts in the application of classifications, which can create a serious amount of bias or systematic error. As a result, the observed patterns in the data may have been artificially produced by factors other than reduction processes, and are wrongly assumed to explain the phenomena being examined (Nance 1987:279). To illustrate this point, consider the practice of removing "usable flakes" from experimental debitage assemblages prior to analysis, as Magne (1985:100, 1989:17) and Ahler (1989a:203, 1989b:99) both do; this is done in order to simulate the removal of useful tool blanks from prehistoric assemblages. However, since the criteria for selecting such flakes are vague, there will be no standardization either between the two types of approaches, or for analysts using only one of them. Consequently, this practice will be unreliable and it will *bias* the flake sample in a systematic manner, each time someone tries to remove one of these flakes.

The issue of *construct* validity is one that applies almost equally to all three analytical methods, since this concerns the fundamental theoretical constructs that guide them. For example, all three assume that lithic reduction occurs predictably in either discrete stages or a continuum, and that reduction techniques such as hard hammer or soft hammer percussion produce distinct patterns of debitage that are readily discernible.

Most researchers tend to overlook a number of weaknesses in these assumptions which can reduce the validity of their approach. For example, since all three of the methods I am studying attempt to differentiate between core reduction and biface manufacture, they assume that certain knapping techniques produce technique-specific flake types or flake class distributions that are capable of identifying the actions that created the assemblage. However, in their experimental evaluation, Ingbar et al (1989) found that neither the Sullivan and Rozen method nor Magne's scar count approach yielded results that were accurate or useful to their research questions. Furthermore, it is known that the specific flake types used by Magne (1985) and sometimes by Ahler (1989a, 1989b), such as the "biface thinning flake" and the "bipolar flake", can be produced by a number of different techniques other than biface thinning and bipolar reduction, respectively; therefore, these named flake types may not serve as useful or accurate indicators of these technologies in the way they are claimed to (Henry et al 1976:57; Sollberger and Patterson 1978:103; Patterson 1982:50; Magne 1985:127; Ahler 1989b: 211).

With regard to reduction *stages*, it has also been found that only the earliest and latest phases of biface manufacture are easy to identify, while flakes generated during the middle of the sequence are not diagnostic and often resemble core reduction debris (Stahle and Dunn 1982; Odell 1989:176; Ahler 1989b:206; Magne 1985:118). In fact, when Mauldin and Amick (1989) tested Magne's approach, they found that dorsal scar counts could not identify reduction stages at all. In addition, when Ingbar et al (1989) used this classification scheme, they found that it was not very valid, probably since Magne's "stages" were based mostly on block core reduction, and therefore might not be applicable to bifacial assemblages like theirs. Since Magne's method and Ahler's Mass Analysis do rely on this stage construct, they may suffer as a result.

B/ RESEARCH DESIGN

Introduction

Researchers involved in lithic analysis emphasize the need for experimentally-derived models that can make our reconstructions of the past more accurate and meaningful. Nevertheless, such models must first be shown to be based on valid and reliable theoretical groundwork. One way to ensure this is by thoroughly testing them, making use of appropriate controls and rigourous experimental design which allows us to systematically investigate the relationships between the variables.

In order for experiments to provide useful information on variability and patterning in the data, the construct validity of the instrument or procedure must therefore be demonstrated. This is accomplished (see Nance 1987:287) by formulating specific predictions or expectations about how the measurements should behave in certain circumstances, followed by the collecting and examining of relevant data, and finally determining whether or not the expectations were met. If they were, then the instrument is considered valid. If not, then either the data were not collected properly, or the theory is wrong, or else the measurement doesn't actually measure the phenomena it claims to. These are the issues that I wish to address in my experimental research.

Experimental Goals

Specifically, I wish to test the following predictions or hypotheses regarding the validity of the three kinds of approaches to debitage analysis used in this study:

Hypothesis 1: The quantification of scar counts on the dorsal and striking platform surfaces of flaking debris will allow the correct classification of all debitage according to its place within a reduction sequence regardless of raw material type, as Magne claims (1985:127).

Hypothesis 2 : The presence of technique-specific debitage, such as biface reduction flakes or bipolar flakes, will accurately identify the method of reduction used to produce a debitage assemblage (see Magne 1985: 115, 126-127).

Hypothesis 3: Debitage is technique-specific, regardless of raw material type. That is, quartzite debitage will show patterning in the flake breakage categories used in the modified Sullivan and Rozen typology, such that it is possible to distinguish core reduction from biface manufacturing debitage, as is claimed. If anything, poorer material will accentuate these flake completeness patterns, not oppose them (Baumler and Downum 1989:107).

Hypothesis 4 : The flake aggregates produced by different technologies and in different stages of manufacture will exhibit markedly different size grade distributions for both obsidian and quartzite (after Ahler 1989b:205).

Experimental Design

Researchers (eg. Jones 1979; Toth 1985; Ahler 1989b) are finding that one successful way to structure their replicative analyses is by performing their experiments with locally available, prehistorically-used raw materials, and by copying a much wider array of technological operations that are thought to be responsible for the variability witnessed in flaking debris. For this reason, I decided to focus my analysis on Morrison quartzite, which was commonly used in the Wyoming region I originally intended to study.

Since I was interested in the morphological traits of quartzite debitage, I developed a double series of replications that included a control set to be carried out with obsidian from Glass Buttes, Oregon, and the other set with Morrison quartzite from near Hyattville, Wyoming. Each of these two sets was to include two examples of the following operations: large prepared block core reduction (each core weighing about 2.5 kg), medium prepared block core reduction (with cores about half the size of the large ones, and weighing roughly 1.3 kg), bipolar reduction, Stage II initial edging of a biface (using a hard percussor) , Stage III primary bifacial thinning (using a soft percussor), and Stage IV secondary bifacial thinning and refining (using soft percussors and pressure flaking). The "Stage" definitions follow Callahan (1979), while I used Prentiss's (1993:84) definition of a prepared core as one in which the platform area has been ground, polished, facetted, or beveled in order to strengthen it; furthermore, the dorsal/platform edge angle of prepared cores ranges from 80 to 90 degrees.

I chose these types of reduction primarily because there has always been a bias toward biface technology in lithic experiments; as a result, we desperately need data from a variety of reduction techniques, such as bipolar work or pressure flaking, in order to make intercomparisons of their products and by-products (Runnels 1985:100; Magne 1989:79).

In order to fully explore the behaviour of quartzite under controlled conditions, and subsequently to assess the validity of approaches designed to analyse flaking debris, I needed to carefully and explicitly structure my experimental procedures so that I could determine which variables had significant effects on the attributes of debitage. The concept of control is particularly important to this aspect of experimental studies, since variables can become confounded; that is, confounding variables may distort experimental results, making it impossible to tell what exactly is influencing the dependent variable (Amick et al 1989:4).

There are two possible ways to deal with the systematic errors produced by confounding variables. One can either randomize the variables being examined, or else the variables can be strictly controlled in order to try and prevent systematic error (Stafford and Stafford 1979:22). I chose to employ methodological rigour and highly controlled conditions surrounding the replications, in order to ensure that any negative evidence is related to a lack of validity in the approaches, and not to some other confounding variable or source of bias (Carmines and Zeller 1979:26). In my experiments, I wanted to control a number of important aspects of the study in order to limit the variability I would see in the debitage to that which was influenced by raw material fracture behaviour.

Throughout these replications, I tried to hold everything constant except for switching the raw material type for half of the sets of reduction operations, so as to avoid introducing any confounding factors or systematic error. The constant variables included the stoneworker (so that I would avoid the bias caused by idiosyncratic knapping behaviours of different people), the raw material type and original size, the knapping tools (i.e. specific hard or soft percussors used consistently), the methods of force application (i.e. freehand percussion, bipolar percussion, or pressure flaking), and the reduction goals (i.e. whether reduction was oriented toward bifacial thinning, or else flake blank production from prepared block cores or bipolar cores). Each reduction "event" was defined as an operation that involves only one type of percussor and one goal, as in

Callahan's (1979) Stage III reduction which can employ a soft percussor to thin a bifacial preform, or as in "hammer-and-anvil" bipolar reduction of a core to produce flake blanks.

The number of blows per event was also held constant for both prepared block and bipolar core reduction, at 20 blows each, whereas, with the bifacial operations, the knapper stopped when he felt it was appropriate to move on to another stage or percussor type. Flakes were collected after each reduction event rather than after each blow was struck, as some researchers have done in the past (eg. Ingbar et al 1989; Magne 1989; Mauldin and Amick 1989; Odell 1989). I did not want to interrupt the flow of the knapper's work in that way, as this could affect the reduction process and potentially introduce a confounding variable.

Finally, since Moloney et al (1988:39) found that control over the flaking process increased with a person's adjustment to the material, I allowed time for the knapper to familiarize himself with the materials, particularly with the quartzite, since I did not want inexperience to affect or bias the debitage produced. This factor has been shown to increase the number of flakes with hinge or step terminations (which would decrease the number of flakes in Sullivan and Rozen's "Complete" category) as well as the total amount of "shatter" (which would affect all three approaches) (Shelley 1990:187; also see Gunn 1975:38). While this might only have been a small confounding variable, I nevertheless wanted to correct for it as much as I could, since it could have created artificial attribute patterning.

In general, my experiments followed this basic procedure:

1) The initial piece of raw material was weighed.

2) The specified knapping operation was carried out by an experienced stoneworker (Dr. Brian Hayden), in the "flaking pit" at the Department of Archaeology, Simon Fraser University. Observations and comments about the operations or the behaviour of the raw material were recorded by me at this time.

3) All flaking debris fell onto a cotton drop-cloth and was collected at the end of each discrete reduction event.

4) The resulting debitage was separated from any remnant cores or finished products and was bagged and labelled according to a constant assemblage number.

5) Debitage was then weighed, sorted, and/or classified according to the general procedures followed first by the modified Sullivan and Rozen typology (see Prentiss 1993), then by Magne's (1985) scar count and attribute analysis of individual flakes, and finally by the Mass Analysis approach (see Ahler 1989a, 1989b). (See below for more details).

6) All data were recorded and entered into a database. The information from the duplicate experimental sets was averaged, to facilitate further quantification and analysis; for example, data from the two quartzite bipolar operations, or from the two obsidian Stage III operations, would be combined and a mean value would then be calculated.

With regard to the three types of debitage approaches mentioned in Step 5, a summary of their analytical procedures is provided below, as I followed these in my analysis.

Modified Sullivan and Rozen Typological Approach: Once the flakes were sorted according to the four size grades added by Prentiss et al (1988), they were classified using the original mutually-exclusive flake categories. These include: 1. "Debris" , or Non-orientable, pieces without evidence of a single ventral surface, a point of applied force, or intact margins; 2. Medial-Distal flake fragments which lack a point of applied force and intact margins but retain a discernible ventral surface; 3. Broken, or Proximal, fragments which retain evidence of a single ventral surface and a point of applied force, but which are lacking intact margins; and finally 4.Complete flakes which show signs of all three attributes mentioned above.

Once the quartzite and obsidian assemblages were categorized using this hierarchical classification key, the percentages of each flake type were compared for each kind of reduction operation employed in the experiments.

Magne's Individual Flake Attribute (Scar Count) Approach: Once the debitage has been collected, those pieces greater than 5mm are retained for analysis, while the rest is ignored. Flaking debris is sorted into two main groups: platform remnant bearing flakes (PRBs) and shatter. Within these two groups, the debitage is examined for flake scars. On platform bearing flakes, the number of scars on the striking platform is counted, whereas, with shatter, the scars on the dorsal surface are counted. If these counts are from 0 to 1, the flakes are classified as "Early"; if the scar count is 2, the flakes are called "Middle", and finally, if there are 3 or more flake scars, the debitage is considered "Late". In addition to these basic classifications, flakes are also examined for features that might label them as biface reduction flakes (which have extensively facetted, acute-angled and often lipped striking platforms) or bipolar reduction flakes (which show evidence of simultaneous percussion from opposite directions, often with crushing) (Magne 1985:100). The data from both the platform remnant bearing flakes and the shatter classes are then pooled into general Early, Middle, and Late categories which are then used to determine the technological origin or place in the reduction sequence.

This method also involves removing possible flake blanks from the debitage collection prior to analysis, although Magne does not provide details on the criteria used when selecting such flakes, except to say that the ones he removed were large (over 30g) and might be suitable for further reduction (Magne 1985:100, 1989:17).

Mass Analysis: Following the separation of flakes from cores and finished products, the flakes were passed through the four different nested screens (G1 to G4) in order to size-grade them. At this point,as in Magne's approach, Ahler (1989b:99) recommends the removal of any flakes or flake fragments that the analyst decides might be useful either as a tool or tool blank , so it is necessary to sort the debitage into two categories: "usable flakes" and "flaking debris". While this practice is admittedly subjective, he considers it to be vital to experimental debitage analysis because it is thought to mimic the same kind of flake removal as was carried out prehistorically. Also, if the analyst wishes, the information on the usable flakes *may* be added back into the main data set for flaking debris, although this is not something that he himself does (see Ahler 1989b: 203).

Nevertheless, once the usable flakes have been separated from the flaking debris assemblage, the analyst then records the weight of each of the four size grades of flakes (***not*** the weight of each individual flake), the total number of flakes in each size grade, and the number of cortical flakes in each size grade, thus yielding a total of twelve variables across the assemblage. This simple data set is then used to calculate new variables such as relative cortex frequency, mean flake weight and so on. Additional information on traditional debitage types such as biface thinning flakes, "shatter", or bipolar flakes can also be added to the data set, although I did not include these extra flake types in my analysis as they did not pertain to the hypotheses I was testing.

Once I had collected the data from these three classification systems, comparisons were made between the known reduction operations, to find out if the different approaches had correctly identified the debitage assemblages. By determining whether or not the theoretical expectations were met for the quartzite flaking debris, it was then possible to assess the validity of the analytical methods.

Expectations

Based on the information gained from petrological research, I suspected that the three methods I was examining would definitely have difficulties correctly classifying quartzite debitage. As mentioned in Chapter Three, this is mainly because the coarser, granular physical structure of quartzite, together with the decreased fracture strength caused by impurities and inhomogeneities in this material, would cause it to fracture differently from obsidian. Consequently, the debitage produced during knapping would be unique enough in its patterns and morphological variability to prevent direct correlation to obsidian or chert debitage distributions. Furthermore, the nature of its physical characteristics may not even allow us to see the important features that define the flake classes used in these debitage approaches for identifying reduction stages or techniques. As Boksenbaum (1980:16) found in his debitage analysis, more of the non-obsidian material was classified as "Uncertain" simply because it was more difficult to interpret its poorer fracture surface.

These suspicions were supported by the few experimental studies which employed non-obsidian and non-chert materials in their replications. From the analysis by Ingbar et al (1989), for example, it was learned that Magne's scar count method and the Sullivan and Rozen typological approach were unable to reliably classify quartzite debitage. These factors may influence the "external validity" of these approaches to debitage analysis; that is, it may not be possible to accurately apply them outside of the original setting that was used to develop the approach, whether this was experimental (using vitreous basalt, in Magne's case) or archaeological (like Sullivan and Rozen's).

In terms of construct validity, these findings, and those of other researchers, may indicate that the theoretical expectations for the classification of quartzite debitage may not be met because of faulty logic in most methods of debitage analysis. For example, some analysts have noted difficulties identifying reduction stages using both Mass Analysis (eg. Patterson 1990:554) and Magne's approach to Individual Attribute Analysis. In fact, Mauldin and Amick (1989:73) tested Magne's (1985) approach and found that there was only a vague trend toward increasing numbers of dorsal scar counts up through the reduction sequence, therefore this method may not be able to identify reduction stages at all. Their suggestion that this method is not valid is echoed by Ingbar et al (1989: 117), who could not find a significant correlation between scar counts and reduction stage, either.

Ingbar et al (1989:132) also noted that the results they produced using the Sullivan and Rozen approach were of little use to them for distinguishing between core and biface reduction techniques. Prentiss and Romanski (1989: 91-92) similarly found that their results contradicted the expectations proposed in the original study.

All of these researchers make vague suggestions about why these validity problems occur and why they may affect our interpretations. Nevertheless, none have yet made an attempt to explore any of these possibilities. Since raw material factors are generally near the top of their lists, it is surprising that my research is the first to systematically explore the effects of different lithic materials on debitage variability. While the theoretical expectations which guide the analytical approaches used in debitage research are based on a certain degree of scientific knowledge, it has become apparent that much remains to be learned about variability in lithic reduction operations and their by-products.

CHAPTER FIVE:
OBSERVATIONS AND RESULTS OF THE ANALYSIS

Introduction

This chapter will discuss the experimental replications and the subsequent debitage analysis, in terms of supporting or negating the hypotheses put forth in Chapter Four. It will explain the four hypotheses and the theoretical expectations associated with them, as well as the results that I obtained from the various methods of analysis. Finally, I will relate these findings to the underlying causes of variability for each hypothesis.

A/ GENERAL COMMENTS ON THE KNAPPING EXPERIMENTS

The knapping experiments proved to be an extremely interesting foray into the world of lithic analysis. Not only did I learn a tremendous amount about debitage and its conditioners, but I also came to a greater understanding and appreciation of the skills needed to work stone materials. Since Dr. Hayden shared his insights into the details and mechanics of knapping both quartzite and obsidian, I was able to add a new dimension to my interpretations of these experimental assemblages.

In general, the obsidian was much easier to knap, partly because of its physical characteristics and partly because Dr. Hayden had more experience working with this lithic material. It clearly shows the features that I was interested in examining, and demonstrates a "best case" scenario in both practical and analytical ways. Nevertheless, the obsidian does have a number of drawbacks which researchers usually do not write about in publications on experimental lithic analysis. To begin with, it is extremely brittle and fragile and, therefore, is somewhat atypical of most stone materials used in the past. Also, despite its ever-favourable reputation, I found that some of the obsidian used in my experiments was of less-than-perfect quality, and sometimes exhibited problematic traits such as gas bubbles (which can decrease the fracture strength) and the presence of phenocrysts or mineral-lined cavities called "vugs" (Shelley 1990:189) which can disturb the even and predictable fracturing of the obsidian. Some of the rejected obsidian was adversely affected by flowbands containing minute crystals that may have been swirled by the flow of the obsidian in its viscous state (Cann 1983:231). Crabtree (1967:17) also noted multiple grades and qualities of obsidian which affected its workability, character, colour, elastic qualities, and so on. He noted that the coarser-textured varieties were weaker and therefore were not as desirable as the more vitreous kinds, like the Glass Buttes material used in my research.

Most researchers agree that its isotropic quality, which enables flakes to be removed from any direction, and its elastic nature make it easy and predictable to work with (Callahan 1979; Crabtree 1967, 1975; Wedel 1978; Luedtke 1992). Furthermore, as Dr. Hayden also noted, less force is required to remove obsidian flakes.

Quartzite, on the other hand, is a much tougher material that often requires tremendously hard and fast follow-through blows to detach flakes from the core or blank (Moloney et al 1988:33). Dr. Hayden found with bipolar reduction, for example, that it was possible to get good flakes with useful edges this way, but that it was very hard to remove them; conversely, with obsidian bipolar work, he only had to tap the core, rather than pound with excessive force. He also noted during block core reduction that it was necessary to use hammerstones that are twice as heavy as would ordinarily be used, and that there are many more unsuccessful attempts at flake removal, compared to obsidian. Effective bifacial (Stage III) thinning of this material was accomplished using a heavy (574g) phenolic resin billet, which is about halfway between wood and antler in terms of hardness (Callahan 1979:168).

Experiments involving pressure flaking operations had to be eliminated, since we found that it was next to impossible to produce invasive flakes on the quartzite in this manner. While it is unfortunate that this technique could not be included in the research design, it may have been unrealistic to ever expect success in this area, since Behm and Faulkner (1974:271) note that pressure flaking of quartzite has been successfully accomplished by very few modern experimenters. Callahan (1979:41) also lists this type of material as beyond the range for pressure flaking since it is not very elastic.

The physical characteristics of the quartzite also affected the reduction process. For example, additional blows were sometimes necessary in order to "normalize" the core and remove irregularities caused by the angular projections of bedding planes. Incipient fracture planes and other inhomogeneities were also troublesome since they can render the material unworkable and unpredictable, thereby making it harder to control the fracture process. Furthermore, on one occasion, an entire block of the quartzite split into several angular pieces as it broke along a series of fracture planes coated with a carbonate lamination. It appears that, before it was recovered from the outcrop area in Wyoming, it had been affected by weathering, perhaps in the form of frost fracturing during the severe winters experienced there and, subsequently, water had soaked into it leaving mineral deposits along the incipient fracture planes. Binford and O'Connell (1984:410) noted the same kind of occurrence in Australia, where weathered surface material had become "rotten" and "broke with a mind of its own", according to their aboriginal informant.

While the cracking open of the Morrison quartzite block was a unique occurrence, it is possible that a similar effect may have taken place on a smaller scale, which may account for the high numbers of angular, flat terminations in my experiments, particularly on bipolar flakes which would ordinarily be expected to exhibit crushing and step fractures on both ends. Other trends in my experimental debitage include the higher numbers of short step and hinge fractures (since some pieces may not have the strength or flexibility to allow detachment of long, thin flakes [Crabtree 1967]), as well as the fewer numbers of flakes produced overall, as compared to obsidian reduction. In the experimental samples analysed here, for example, the total number of quartzite flakes was 5152 versus 7290 for identical operations using obsidian. However, the effects of raw material composition and physical properties will be discussed in more detail throughout the remainder of this chapter.

EVALUATION OF HYPOTHESIS # 1

Hypothesis 1: The quantification of scar counts on the dorsal and striking platform surfaces of flaking debris will allow correct classification of all debitage according to its place within a reduction sequence regardless of raw material type, as Magne (1985:127) claims.

Expectations

Despite Magne's (1985:120) argument that the ordinal classification of all types of debitage can identify general reduction stages, the results of my analysis have shown that this is not the case. To arrive at this conclusion, however, and to assess the validity of his method, it was first necessary to find out what the specific theoretical expectations were that guided its methodology.

To begin with, Magne (1985) assumes that two debitage variables (dorsal scar count and platform scar count) are capable of reconstructing stone tool manufacturing stages. Since the number of flake scars on the dorsal and striking platform surfaces is expected to increase as lithic reduction proceeds (Magne 1985:113-114), he has suggested that certain numbers of scars be equated with certain reduction stages. Specifically, he argues that flakes bearing 0 to 1 scars are indicative of "Early" stage events including single platform, and bipolar, core reduction. Debitage with 2 flake scars, however, are classified as "Middle" stage reduction debris, which he argues is produced during the primary trimming stages of tools, including reduction of marginal retouch tools (i.e. flake tools such as scrapers) and the first half of the reduction events of all other tools, whether unifacial (such as endscrapers) or bifacial. In other words, this stage is said to be mainly related to edge straightening and the removal of most of the excessive mass of a tool blank. Finally, any flakes bearing three or more scars are considered to represent the latter half of reduction events of unifacial and bifacial implements, which is related to refining the intended shape of the tool. As such, these flakes are classified as "Late" stage debris (Magne 1985:106-107).

In essence, then, a person using this classification should expect to be able to separate different reduction technologies like hard hammer core reduction from soft hammer percussion, as well as be able to separate the various stages of tool manufacturing sequences. Magne (1985:127) argues that this success will not be affected by differences in raw material characteristics.

This claim is somewhat misleading, however, since, in his own experiments, Magne (1985:127) found that certain raw material types had advantages over other types; for example, obsidian was found to have a slight discriminatory advantage over basalt due to the fact that flake scars were more easily observable on its glassy black surface. Despite this apparent contradiction, Magne (1985:127) considers such differences to be "a sort of systematic error factor" that can basically be overlooked. Subsequently, he goes on to claim that the differences between raw material types in stage variability are not very great, a fact which he argues was supported by his experimental findings.

To summarize, then, analysts using this method could expect that debitage of any and all raw material types will bear standard numbers of flake scars that directly correspond to reduction technologies or manufacturing stages. They could expect core reduction assemblages to be dominated by debitage with no flake scars or one flake scar on the dorsal surface or on the striking platform. Furthermore, edging and thinning debitage is predicted to have 2 flake scars on these areas. However, if flakes bear three or more scars, then it should be safe to assume that the assemblage was produced by late stage tool refining procedures such as billet or pressure flaking.

If this method is indeed a *valid* one, then all of these theoretical expectations should be met, whether the debitage is made of obsidian, basalt, or quartzite.

My Experimental Results

Hypothesis 1: Negated

Contrary to Magne's claim (1985:127), the quantification of flake scars on the dorsal and striking platform surfaces of flaking debris does not allow correct classification of all debitage according to its technological origin or reduction stage, regardless of raw material type.

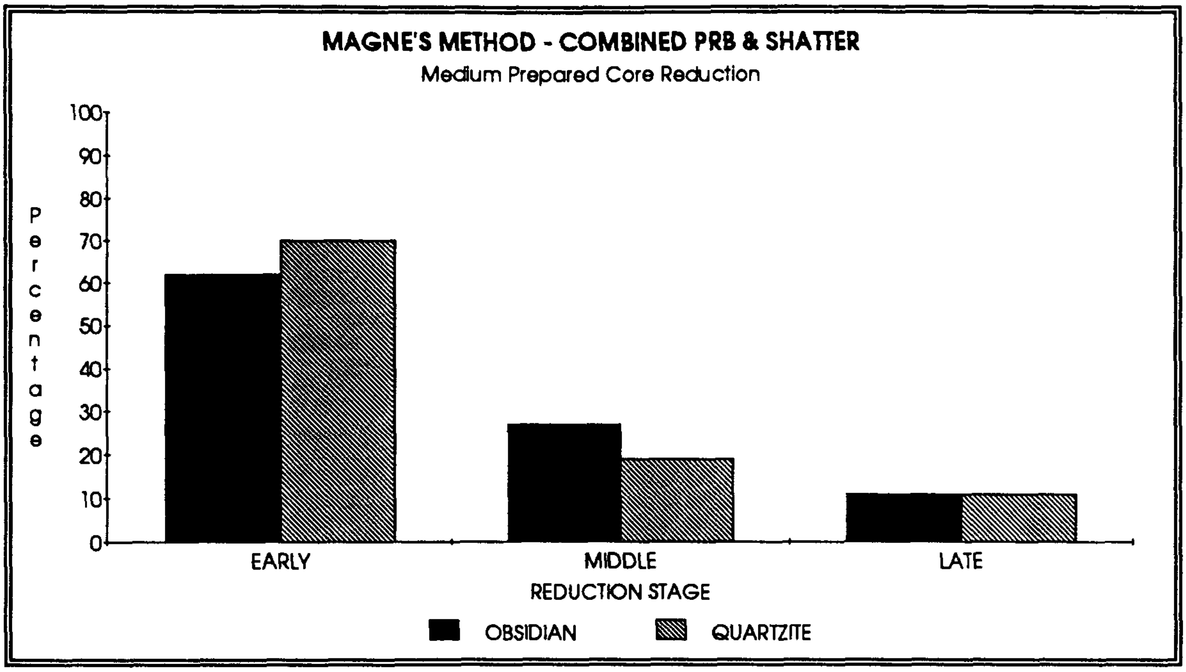
Like Ingbar et al (1989), I found that Magne's classification is not very accurate. While the "Early" debitage class did have more flakes than either of the other two classes in the assemblages produced by medium and large prepared block core reduction as well as bipolar operations, as Magne predicted it would, his method was unable to correctly classify the flaking debris from Stage II bifacial reduction (which is equal to his "Middle" stage) or the quartzite flakes from Stage III or soft hammer percussion (his "Late" stage). Essentially, almost all of the assemblages were dominated by so-called Early stage debris, regardless of the technology used to produce it, or its place in a reduction trajectory. The validity of his expectations will be explored individually, below.

a) Is "Early" stage debitage produced by core reduction, as expected?

While the three types of core reduction employed in my experiments (medium prepared core, large prepared core, and bipolar) did produce assemblages dominated by "early" debitage with 0 or 1 flake scars, this same phenomenon also occurs for all the other kinds of reduction operations as well, with the single exception of the obsidian Stage III soft hammer sample which is dominated by "late" stage flaking debris. Therefore, this method of analysis is unsuccessful in the first of its goals, since, for the most part, it failed to segregate core reduction from other technologies.

With the medium-sized prepared block cores (see Figure 1), almost half of the quartzite and obsidian assemblages are made up of "early" shatter with 0 or 1 dorsal flake scars. This occurrence is fairly consistent with quartzite across all types of reduction, whereas this is an anomaly for obsidian. Large prepared block cores are somewhat different, since such high numbers of "early" shatter are only found in quartzite samples (see Figure 1); obsidian, on the other hand, has less of a disproportionate distribution between "early" and "middle" debris. Platform-bearing flakes contribute much less to these assemblages, and the "early" examples of these flakes contribute less than half the proportions of flakes as the "early" *shatter* does (averaging about 18%). In both of these types of prepared core operations, quartzite always produces more flakes than obsidian does.

In terms of bipolar core reduction, the quartzite assemblage is vastly different from the obsidian (see Figure 2). With quartzite, almost 70% of the assemblage is composed of "early" shatter, while the same category in the obsidian sample is much less, at only 44%. Each of the three obsidian shatter classes comprise about a third of the collection, giving it a very even distribution that is quite unique among the experiments.



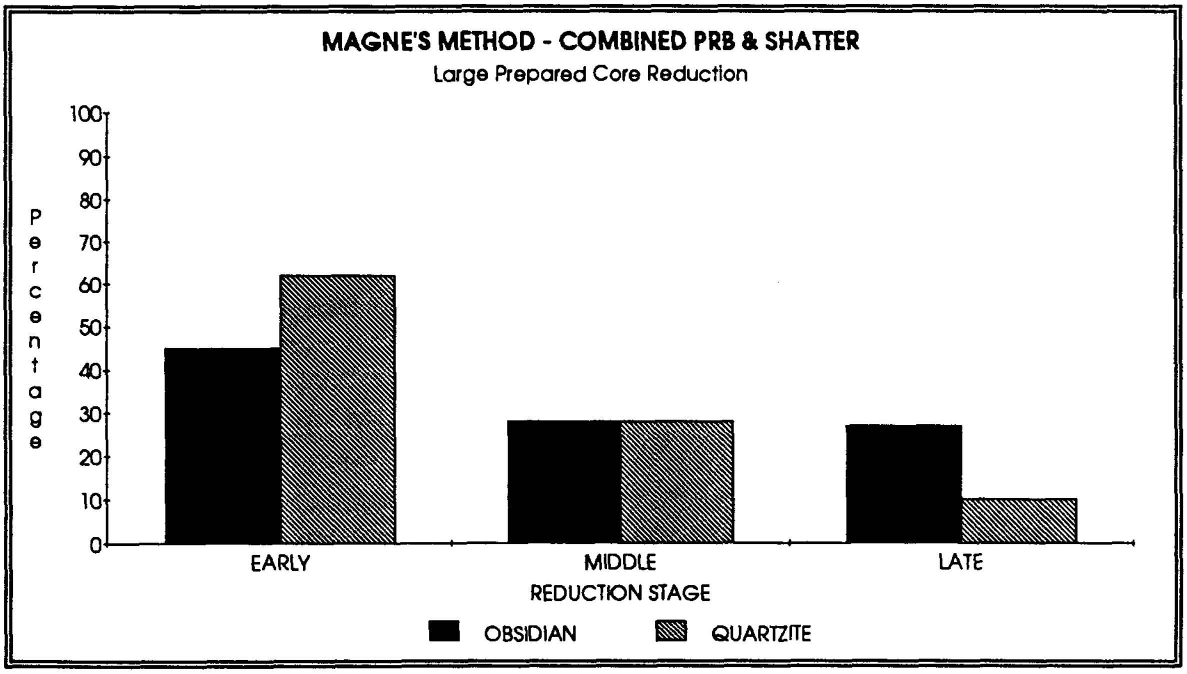
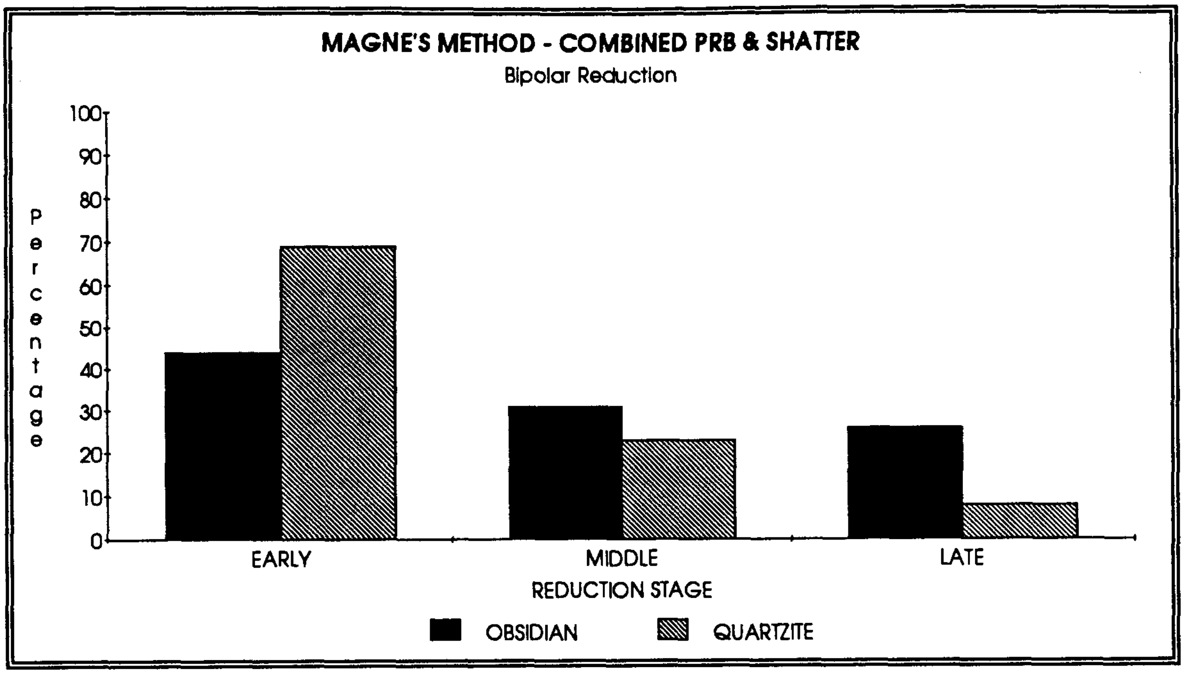
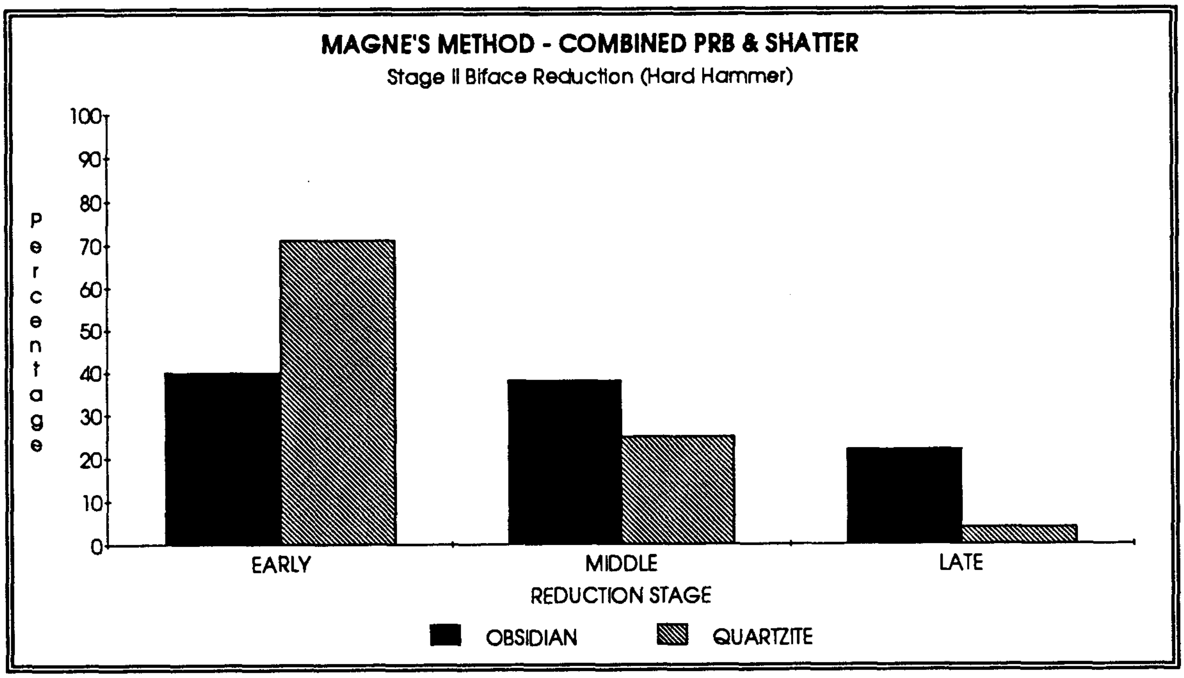


Figure 1: Identification of the reduction stages in experimental Medium and Large Prepared Core debitage, according to Magne's scar count method. Counts of platform remnant bearing flakes (PRBs) and shatter are combined.
("Early" = 0-1 flake scars; "Middle" =2 flake scars; "Late" = 3 or more flake scars)





<u>Figure 2</u>: Identification of the reduction stages in experimental Bipolar and Stage II Biface debitage, according to Magne's scar count method. Counts of platform remnant bearing flakes (PRBs) and shatter are combined.
("Early" = 0-1 flake scars; "Middle" = 2 flake scars; "Late" = 3 or more flake scars)

b) Is Magne's "Middle" stage reduction debris produced by primary thinning and edging of tools (what I called Stage II reduction) ?

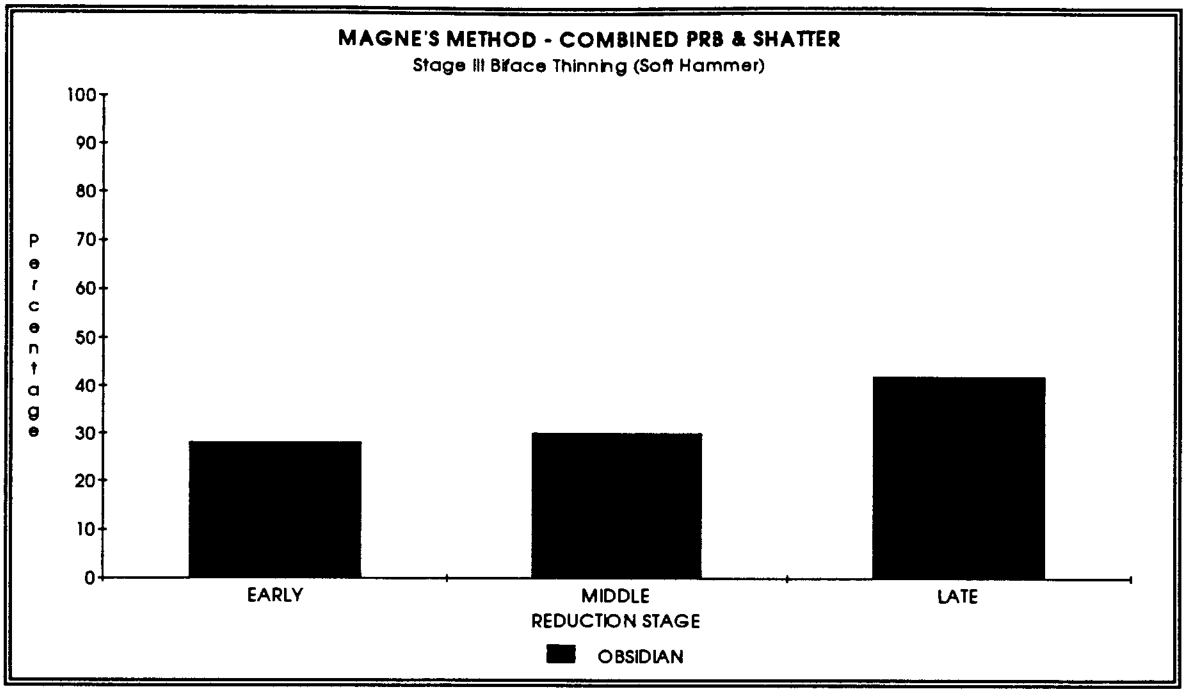
In all of the block core assemblages, "middle" stage shatter is fairly common, comprising an average of 26% across all operations. However, middle stage debris is said to represent primary edging and thinning of tools, so it is surprising that its proportions are so high in core reduction assemblages. In this sense, Magne's method is not very useful, since it misidentifies a fairly large portion of core reduction debris as if it were primary thinning debris.

Furthermore, in the actual Stage II assemblages (see Figure 2), I found that there is *less* "middle" stage reduction debris (at 37%) in the obsidian sample than the "early", or "core reduction", class (at 40%). With quartzite, as well, the majority (71%) of the Stage II flaking debris is classified as *core* ("early") reduction, and only 25% is actually correctly identified as "middle" stage debris. This phenomenon indicates that Magne's method is not able to produce valid or accurate results in this case, either.

c) Is Magne's "Late" stage debitage produced by soft hammer percussion?

The composition of the "late" stage, or Stage III, reduction assemblages best highlights the potential problems related to extrapolating results from one type of lithic raw material to that of another, since this kind of reduction operation is the only one that behaved as Magne predicted, but only for the *obsidian* debitage (see Figure 3).

In this case, I found that much (42%) of the obsidian debris is classified as "late" stage material, thereby providing a relatively accurate representation of the knapping operation that had taken place. In this case only, Magne's method is successful. Nevertheless, when the results of the quartzite classification are examined, the exact opposite is found. Once again, the quartzite debitage is dominated by "early" stage, or core reduction, flakes,



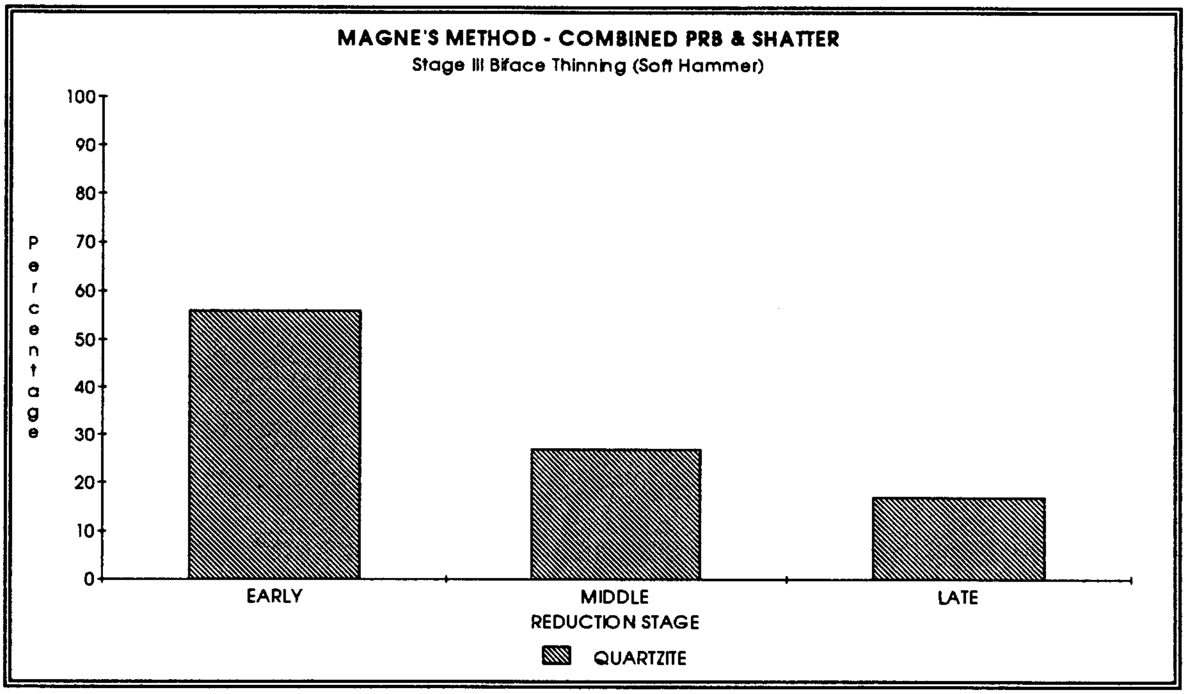


Figure 3: Identification of the reduction stages in experimental Stage III (Soft Percussor) Biface Thinning debitage, according to Magne's scar count method. Counts of platform remnant bearing flakes (PRBs) and shatter are combined.
("Early" = 0-1 flake scars; "Middle" = 2 flake scars; "Late" = 3 or more flake scars)

which make up 58% of the collection. In addition, the "late" flakes that should have been in the majority comprise only 15% of the assemblage.

How the inaccuracy and invalidity of the Individual Flake Attribute Approach is related to raw material factors

Since the criterion for classifying debitage is the highest number of flake scars either on the dorsal surface (for non-platform-bearing flakes) or the striking platform (for platform-bearing flakes), there are several important petrological factors that will influence both the formation of such flake scars and the ability to observe them clearly.

In other research, including Magne's (1985) own, coarser-textured materials are renowned for obscuring important fracture attributes used to determine reduction strategies (see Patterson 1990:550-551). For example, in the study by Ingbar et al (1989:123), they report that quartzite was correctly classified according to reduction stage less often than microcrystalline chert. They found that, with the dorsal scar counts, Morrison quartzite achieved only 40% classification accuracy for shatter, almost 17% less than that for chert shatter; in addition, with *platform* scar counts, Morrison quartzite debitage was accurately classified only 35% of the time, compared to 48% for chert flaking debris. I also found decreased accuracy in the classification of quartzite as compared to obsidian.

Although most of the quartzite "shatter" flakes have fewer than two or three observable dorsal flake scars, regardless of the reduction strategy used, it is possible that the coarser texture of this material is masking the ridges formed by additional flake scars; however, it is more likely that quartzite merely fractures differently and produces fewer flakes in general (see Moloney et al 1988), because it is not as brittle as obsidian or other vitreous materials like the basalt that Magne also based his conclusions on. Since shatter is thought to be caused by marginal fracturing of striking platforms, bending of feather terminations, or platform trimming (Shafer 1985:294-295), these actions may produce lower percentages of quartzite shatter, compared to obsidian. Consequently, this would affect Magne's platform-bearing flake:shatter ratio, which is a measurement that he says may be indicative of gross reduction stage (Magne 1989:17; see also Magne 1985:104-106).

While his calculations of this ratio were only intended to be a preliminary estimate of technological diversity, he noted a few possible trends in the data, which I have included together with the ratios from my own experiments, where applicable (see Table 1). I found that obsidian and quartzite behaved in very different and unpredictable ways. For example, in comparison to obsidian, quartzite bipolar and medium-sized core operations produced more platform flakes than shatter but, in the other hard hammer operations, the ratios were *lower* for quartzite. In addition, in debitage derived from soft hammer biface reduction, quartzite produces over 50% more platform flakes than obsidian does; this is the most striking difference in all the experiments.

In terms of Magne's results, the only similarities are in the very low relative values for bipolar reduction and the elevated ratios in the bifacial assemblages. However, since Magne did not intend these values to be direct indicators of reduction technique, this information is included here mainly as a preliminary measure of raw material and, to a lesser degree, technological differences that can be manifested in debitage collections.

With regard to platform remnant bearing flakes (PRBs), I encountered two interesting phenomena that severely affect the number of flakes that could be placed in the three PRB classes. The first is the presence of "split flakes" which generally are characterized by orthogonal fractures that bisect the striking platform and axis of percussion of some lithic flakes (see Figure 4a). These have also been referred to as "Siret accidents" (Moloney et al 1988:32) or "pseudo-burins de Siret" (Bordes 1961). These flakes are thought to be produced when heavy percussors are used in the knapping

PRB/SHATTER RATIO

	OBSIDIAN	QUARTZITE	MAGNE'S BASALT
Bipolar	.03:1	.02:1	.15:1
Bipolar	.01:1	.13:1	
Med. Core	.28:1	.40:1	
Med. Core	.26:1	.24:1	.83:1
Large Core	.15:1	.29:1	
Large Core	.62:1	.26:1	
Stage 2	.32:1	.27:1	
Stage 2	.48:1	.20:1	
Stage 3	.37:1	.54:1	.45:1
Stage 3	.23:1	.78:1	
Averages:			
Bipolar	.02:1	.08:1	
Med. Core	.27:1	.32:1	
Large Core	.39:1	.28:1	
Stage 2	.40:1	.24:1	
Stage 3	.30:1	.66:1	

Table 1: Ratio of Platform Remnant Bearing flakes to Shatter, for my experimental obsidian and quartzite debitage and Magne's (1985) experimental basalt debitage.

process; consequently, the extra force sometimes exceeds the fracture strength or elastic capabilities of the material and the flake splits right down the middle or occasionally on an oblique angle (see Figure 4b) (Boksenbaum 1980). In my experiments (see Table 2) they occur most often in bipolar and block core reductions, and somewhat less often in Stage II (hard hammer) biface thinning, ranging on average from 11 to 15% of the assemblage for quartzite, and 6 to 8% for obsidian. They are the least common in billet flaking operations, particularly those involving obsidian, where the average percentage of such flakes is only 2% of the total flake count.

It is my suspicion that the higher frequency of split flakes in quartzite flaking debris has to do with certain physical properties of this material, especially toughness, which requires the knapper to use larger, heavier percussors and more intense blows in order to remove flakes. Since obsidian knapping requires much less force and lighter percussors for successful flake removals, it is not surprising that fewer split flakes are encountered in its debitage collections.

The occurrence of "secondary multiple flakes" is another phenomenon that is rarely discussed in lithic analysis; as such, it was hard to find information on how they are caused, and next to impossible to determine in what frequencies they have occurred in various archaeological or experimental lithic reduction technologies. These types of flakes are produced simultaneously with the primary flake, and show superimposed positive and negative bulbs of percussion (Jelinek et al 1971:198) (see Figure 5). They are easy to identify because the central dorsal scar and the ventral surface have the same point of applied force (Boksenbaum 1980:13); this creates a thinning or shearing of the striking platform which gives it an unusual curved morphology that is often shaped like an S or a rounded V rather than the usual lens shape. These have been referred to by some archaeologists as "seagulls" (Rahemtulla pers. comm. 1994) because, from above, the platforms can look like the schematic representation of this bird in flight.

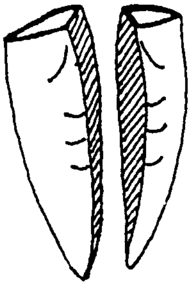


Figure 4a: Split Flake showing orthogonal fracture.

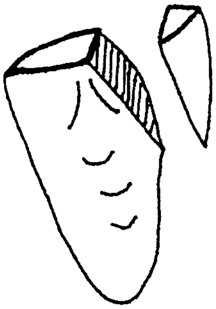


Figure 4b: Split Flake showing oblique fracture.

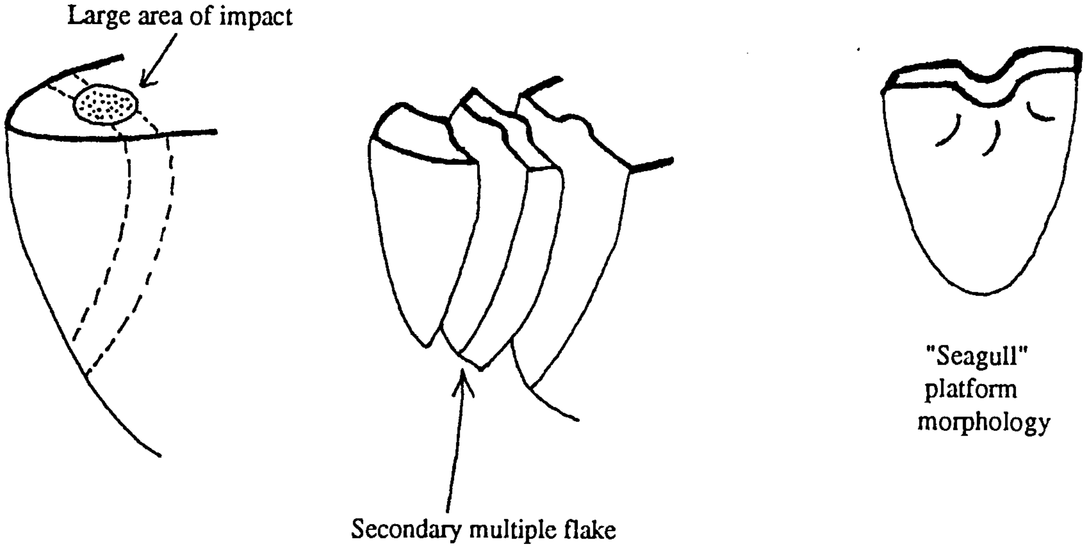


Figure 5: Secondary Multiple Flake.

Frequency of Split Flakes (as a % of total flake counts/assemblage)

	OBSIDIAN	QUARTZITE
Bipolar	6	21
Bipolar	7	2
Med. Core	3	13
Med. Core	3	14
Large Core	1	11
Large Core	7	20
Stage 2	1	5
Stage 2	2	18
Stage 3	5	9
Stage 3	2	6
Averages:		
Bipolar	7	11
Med. Core	2	13
Large Core	4	16
Stage 2	2	11
Stage 3	4	7
Overall averages:	4%	12%

TABLE 2: Occurrence of Split Flakes expressed as a percentage of the total debitage count for obsidian and quartzite experimental knapping operations.

Secondary multiple flakes are thought to be produced when there is a broad area of contact between a large, blunt hammerstone and the core, which may knock off two or more flakes at the same time. In my experiments, I found that they were virtually absent in all bipolar assemblages, but they were a fairly common occurrence in the other hard hammer operations on both obsidian and quartzite, especially block core reduction (see Table 3). This is consistent with the suggestion that they are generally associated with flat or convex platform surfaces like those found on most cores, rather than with concave surfaces like those present on bifacially-thinned items (Jelinek et al 1971:200), although I did encounter a few of these flakes in soft hammer operations.

In general, they are more common in obsidian assemblages, perhaps because this material is much more elastic than quartzite and therefore it will not break when the material fractures in this way; the coarse nature and decreased structural integrity of the quartzite may cause this material to shatter more often under the severe stress of intense hard hammer percussion. These flakes are predominantly found in the smaller sized flake-grades (Ahler's G3) and in Magne's "Early" categories, however, they still only make up an average of 6% and 11% of the entire debitage assemblages for quartzite and obsidian, respectively.

Nevertheless, both secondary multiple flakes and split flakes have a serious effect on Magne's flake classification procedure because one of his steps relies on the quantification of flake scars on the striking platform. If the striking platform is not all there, either due to orthogonal fractures that remove one half of it, or due to secondary detachments that can shear it down even more, then it is not possible to count its flake scars. Consequently, these flakes (which are usually otherwise complete) must be classified as "shatter", even though part of the platform can still remain. Magne (1985: 114) himself noted the difficulties in dealing with narrow platforms such as these, and noted that those less than 2mm wide were usually considered as shatter.

Frequency of Secondary Multiple Flakes (as a % of total flake counts/assemblage)

	OBSIDIAN	QUARTZITE
Bipolar	4	5
Bipolar	0	<1
Med. Core	16	13
Med. Core	21	9
Large Core	14	8
Large Core	22	6
Stage 2	15	6
Stage 2	9	10
Stage 3	8	8
Stage 3	2	4

Averages:

	OBSIDIAN	QUARTZITE
Bipolar	19	11
Med. Core	18	7
Large Core	12	8
Stage 2	5	6
Stage 3	2	3
Overall Average:	11%	6%

By Reduction Stages:

	OBSIDIAN	QUARTZITE
Early	6	4
Middle	3	1
Late	2	1

Table 3: Occurrence of Secondary Multiple Flakes expressed as a percentage of the total debitage count for obsidian and quartzite experimental operations.

More importantly, shearing produced by secondary flake detachments will also create additional flake scars on the dorsal surface which are unrelated to sequential thinning in the way that Magne claims. Consequently, many of these flakes will have to be classified as "Middle" or "Late" stage shatter, by definition (since they have two or more dorsal scars), even though the knapping operations that created them may not have had anything to do with such reduction activities. For example, core reduction or early stage biface edging debitage that has had extra dorsal flake scars created by the removal of secondary multiple flakes will be classified as "Late" stage flaking debris since the scar count is more than two; therefore, an analyst using Magne's approach could inaccurately interpret this assemblage as one produced by billet flaking or final edge straightening and thinning operations. Unfortunately, this can create a serious systematic error in the classification procedure which can, in turn, create artificial patterning in the data.

I observed this phenomenon over and over again in my research. For example, in Assemblage 16 (a Stage III soft hammer percussion assemblage), one third of the "Late" shatter was made up of secondary multiple flakes with sheared platforms which, judging from the platform *remnants*, would otherwise be classified as Early Stage PRBs. However, because of Magne's classification criteria, I had to place these flakes in the "Late" shatter category because the number of dorsal flake scars was greater than two.

Therefore, the main theoretical flaws in Magne's approach are that increased numbers of dorsal flake scars are not always caused by progressive thinning operations, nor do flake scars identify technological origin regardless of raw material differences, as Magne incorrectly assumes.

EVALUATION OF HYPOTHESIS # 2

Hypothesis 2: The presence of technique-specific debitage, such as biface reduction flakes or bipolar flakes, will accurately identify the method of reduction used to produce a debitage assemblage (see Magne 1985:115, 126-127).

Expectations

Analysts using Magne's approach are provided with definitions of biface reduction flakes and bipolar flakes which he says should allow identification of these technique-specific flake types. Biface reduction flakes are said to have extensively facetted striking platforms which are acutely angled and often "lipped" (Magne 1985:100). These are claimed to be associated with his "Late" stage reduction operations and therefore they can be expected to identify the occurrence of billet flaking in a debitage assemblage.

Flakes are classified as bipolar flakes, on the other hand, if they show evidence of simultaneous percussion from opposite directions, often with crushing (Magne 1985:100). An analyst who encounters such flakes in his or her collection can expect this presence to indicate that the debitage was formed by the bipolar, or hammer and anvil, reduction technique.

One possible problem that could be encountered is the fact that "flake types, no matter carefully defined, are not linked on a one-to-one basis with particular technologies" (Ahler 1989a:211). Consequently, I could expect that flakes with similar attributes could be produced by any force application method (Patterson 1982:50). Hayden and Hutchings (1989:247), for example, found that lipping can occur with *hard* hammers (although rarely), and Ahler (1989a:211) noted that a small number of so-called biface thinning flakes and bipolar flakes were produced by hard hammer bifacial edging operations (what I refer to as "Stage II" reduction). In addition, in Magne's experiments, he found that bipolar reduction could produce biface reduction flakes (1985:127). Therefore, while such rare occurrences may not affect overall relationships, it was safe to assume that I could observe similar discrepancies between the reduction technology and the specific flake types.

Results of My Experiments

Hypothesis 2: Negated

Bifacial reduction flakes and bipolar reduction flakes are not excellent indicators of each type of reduction, as Magne (1985: 115,126) claims. Neither are they always accurately identified by the same variables used to predict early, middle, and late reduction stages (Magne 1985:115).

a) Biface Reduction Flakes

Magne (1985:100) defines biface reduction flakes (BRFs) as those exhibiting "extensively facetted, narrow angle and often 'lipped' platforms". In my experiments I found that the relative frequency of biface reduction flakes is very low in both quartzite and obsidian assemblages, each averaging only 5% of the Stage III assemblages (see Table 4). Unfortunately, this supports the claims of Magne's *opponents*, rather than the expectations proposed by Magne himself. Furthermore, biface reduction flakes are encountered in quartzite large block core operations as well as in quartzite and obsidian medium core reduction. Consequently, this shows that biface reduction flakes do not serve as good indicators of late stage bifacial reduction, since they occur in relatively inconsequential numbers in both obsidian and quartzite Stage III biface thinning operations and also since they can be found in hard hammer core reductions, both contrary to Magne's expectations.

OBSIDIAN

	BRFs (%)	Late PRBs (%)	Combined BRFs and Late PRBs (%)
Bipolar	0	1	1
Bipolar	0	0	0
Med. Core	1	2	3
Med. Core	0	2	2
Large Core	0	0	0
Large Core	0	23	23
Stage 2	0	5	5
Stage 2	0	6	6
Stage 3	5	9	14
Stage 3	6	4	10
Averages:			
Bipolar	0	1	1
Med. Core	1	2	3
Large Core	1	14	15
Stage 2	0	6	6
Stage 3	5	7	12

QUARTZITE

	BRFs (%)	Late PRBs (%)	Combined BRFs and Late PRBs (%)
Bipolar	0	0	0
Bipolar	0	1	1
Med. Core	0	0	0
Med. Core	1	0	1
Large Core	1	2	3
Large Core	0	1	1
Stage 2	0	0	0
Stage 2	0	1	1
Stage 3	1	1	2
Stage3	7	7	14
Averages:			
Bipolar	0	1	1
Med. Core	1	0	1
Large Core	1	2	3
Stage 2	0	1	1
Stage 3	5	4	9

TABLE 4: Occurrence of Biface Reduction Flakes and Late Stage Platform Remnant Bearing Flakes in obsidian and quartzite debitage assemblages, according to Magne's scar count method.

b) Bipolar Reduction Flakes

Once again, contrary to the expectations I had going in to the experiments, the number of bipolar flakes that I actually encountered was very low in the bipolar assemblages (see Table 5). With obsidian, bipolar flakes, which Magne (1985:100) says are identified by "evidence of simultaneous percussion from opposite directions, often with crushing", make up only 9% of the total flake count for this type of operation, and this amount is even less for the quartzite bipolar debris, comprising a mere 6% of the collection. Furthermore, one "bipolar flake" was produced during Stage III soft percussor thinning of an obsidian biface.

While the single bipolar flake found in "Late" stage reduction might be considered an unimportant anomaly, the low relative frequencies of this flake type, as well as the low biface reduction flake frequencies, do demonstrate a vague trend in the flaking debris, since, for the most part, these flake types do not occur in other types of reduction operations. Unfortunately, this trend negates Magne's (1985: 126) claim that such named flake classes act as "excellent indicators" of these types of reduction.

How raw material factors are related to the failure of named flake types for indicating reduction techniques

One possible reason that Magne's biface reduction flake and bipolar flake categories do not work well as indicators of technological strategies is the narrowness of the definitions he provides (see Magne 1985:100). Since his criteria are rather limited and vague, they do not cover the range of variation that can be found among even specific flake types like these; as a result, I found that it was sometimes difficult to tell whether a particular piece of debitage should be placed in one of these specific categories. While the attributes used to identify biface reduction flakes were quite basic and standard, his definition of bipolar flakes was very sketchy and it greatly limited the number of flakes

BIPOLAR FLAKES (%)

	OBSIDIAN	QUARTZITE
Bipolar	3	8
Bipolar	20	3
Med. Core	0	0
Med. Core	0	0
Large Core	0	0
Large Core	0	0
Stage 2	0	0
Stage 2	0	0
Stage 3	0	0
Stage 3	2	0
Averages:		
Bipolar	9	6
Med. Core	0	0
Large Core	0	0
Stage 2	0	0
Stage 3	1	0

TABLE 5: Occurrence of Bipolar Flakes in experimental obsidian and quartzite debitage, expressed as a percentage of the total flake count per assemblage, according to Magne's scar count method.

that I would classify as bipolar flakes. This occurred because he never explicitly defined what he considered to be the evidence for bipolar reduction in the way that Kobayashi (1975) has, for example. Perhaps he should have provided a more thorough, or polythetic, list of attributes that would clearly have identified this type of flake. Hayden (1980:3; Hayden and Hutchings 1989) has found such an approach to be the most reliable way of classifying both bipolar and billet flakes.

Nevertheless, raw material factors also appear to play a very important role in the negation of this hypothesis. For example, with obsidian, which is so brittle, the striking platforms tend to shatter to such a degree that they either disappear altogether (causing the flakes to be classified as "Late" shatter) or else so much of them is missing that an analyst cannot determine the number of flake scars on the platform surface. Consequently, it may not be possible to observe the facetting or lipping that is thought to identify biface reduction flakes.

With quartzite, as mentioned earlier, striking platforms tend to split both down the axis of percussion or on an oblique angle, taking off the corner of a flake (see Boksenbaum 1980:14), or else several flakes become detached at the same time, thereby removing all or most of the striking platform. Again, this makes it impossible for an analyst to quantify flake scars on this feature. The coarse texture or grain size will also detrimentally affect the visibility of such flake scars. Unfortunately, we must accept the fact that "the mechanisms responsible for flake formation will not always be apparent" with some lithic raw materials (Cotterell and Kamminga 1987:704).

In general, Morrison quartzite just does not allow as great a degree of thinning as obsidian does, therefore facetting will not always be present in the way that Magne claims. Examination of the finished Stage III bifaces shows that the greater edge thickness of the quartzite bifaces does not allow for the production of numerous narrow-angled biface reduction flakes in the way that obsidian does. Evidence to support this is found in the relatively low frequency of "Late" stage platform-remnant bearing flakes (including biface reduction flakes) in quartzite billet flaking assemblages, as compared to obsidian.

With regard to bipolar debitage, raw material factors are important as well. While obsidian is more likely to demonstrate the crushing that Magne uses to define bipolar flakes (since it is so brittle and vitreous), quartzite is much less likely to fracture in this expected manner. I found, with the Morrison quartzite, that most flakes derived from bipolar reduction quite often have completely *flat* terminations and initiations, as opposed to crushed or sheared fracturing. I believe that this has to do with the numerous incipient fracture planes and bedding planes frequently found in this material (see Duncan 1969); when undergoing bipolar percussion, the "wedging" fracture (see Cotterell and Kamminga 1987) at both ends may have caused the shattered areas to break off in larger, angular chunks along these planes, rather than retaining the crushed appearance that is preserved in obsidian bipolar flakes.

Because of this petrological characteristic, the number of flakes classed as bipolar flakes in the quartzite assemblages was greatly reduced. Consequently, Magne's classification failed to correctly correlate a named flake type with its technological origin, just as it did with biface reduction flakes. The failure of this approach demonstrates yet another way that raw material factors adversely affect the validity of the methods and criteria used in approaches to debitage analysis.

EVALUATION OF HYPOTHESIS #3

Hypothesis 3: Debitage is technique-specific, regardless of raw material type. That is, quartzite debitage will show patterning in the flake breakage categories used in the modified Sullivan and Rozen typology, such that it is possible to distinguish core reduction from biface manufacturing debris, as is claimed. If anything, poorer material will accentuate these flake completeness patterns, not oppose them (Baumler and Downum 1989: 107).

Expectations

The original Sullivan and Rozen typology (1985) was designed to distinguish between biface production and core reduction on the basis on flake completeness data. They expected that core reduction debitage would be dominated by high numbers of complete flakes and nonorientable debris, whereas biface production, or "secondary reduction", would have high numbers of broken ("proximal") flakes and medial-distal fragments produced when thin billet flakes break (Sullivan and Rozen 1985: 762, 769, 682).

However, Prentiss and Romanski (1989:92) found that their experimental chert debitage did not live up to these expectations. Instead, they found that biface manufacture produced far more complete flakes than core reduction, and that proximal flakes were almost as common in both tool and core reduction. In core reduction, they found that there were great quantities of nonorientable debris and only moderate numbers of complete flakes and medial-distal fragments, contrary to the original expectations put forth by Sullivan and Rozen. Furthermore, split flakes were not found only in the initial phases of core reduction (Sullivan 1987), but were produced in low numbers in both kinds of reduction technologies.

Even though three experimental assessments of the Sullivan and Rozen approach found it to be somewhat successful at distinguishing between core and biface reduction (Baumler and Downum 1989:113; Ingbar et al 1989:121; Prentiss and Romanski 1989:91), I suspected that there would be some problems with the debitage patterns being inaccurate for characterizing quartzite operations. This is based on the assumption that some lithic materials, which are different petrologically, will probably create unexpected debitage profiles. Brisland (1992:99), for example, found that obsidian consistently produced greater proportions of medial-distal fragments than either chert or basalt. In addition, Prentiss and Romanski (1989:93) cautioned that the fracture properties of different raw materials would greatly affect flake breakage and the subsequent debitage patterning. I would expect that a more coarse-grained material like quartzite would demonstrate even more marked differences across the flake-breakage categories.

In terms of the modified version of this approach (the MSRT), the size grades added to the basic flake classes seem to enable the recognition of more distinct patterns in the debitage depending upon the reduction technology used (Prentiss 1993:133). Therefore, going into my experiments, I expected that the MSRT would eliminate most of the problematic aspects noted by the previous researchers, with the exception that the quartzite patterns might not be quite as clear as those for my obsidian operations.

I expected to find high numbers of complete flakes and equal numbers of proximal and split flakes in biface production operations (my Stage III biface thinning), and very large amounts of nonorientable debris, together with moderate numbers of complete flakes and medial-distal fragments, in the prepared core reductions. In terms of the four size grades, I expected that the extra-large and large flakes would predominantly be complete, and that they would mostly be associated with hard hammer core reduction; medium and small sized flakes, however, were presumed to occur in both core and biface assemblages, primarily as a result of platform preparation and breakage in the larger categories (Prentiss 1993:130-132).

Of all the researchers who have employed this approach, only Prentiss (1993) has commented on how I could expect the debitage from bipolar reduction to be patterned across the flake classes and the additional size grades. Therefore, I examined this unique flaking technique as an additional sideline of research that could be used to test these expectations and to expand our understanding of the variability inherent in debitage from this kind of reduction. In general, Prentiss (1993:431) said to expect that there would be high numbers of small fragmentary debitage (that is, the nonorientable and medial-distal classes) because the intense crushing and wedging actions associated with the hammer and anvil technique are likely to eliminate most complete or proximal flakes.

Results of My Experiments

Hypothesis 3: Negated

The modified Sullivan and Rozen typology (MSRT) failed to clearly differentiate between prepared core reduction and biface production. For the most part, these two technologies did not develop "largely independent patterning" in the various size categories, as Prentiss (1993:236) proposed. Furthermore, contrary to Baumler and Downum's (1989:107) claim, the "poorer" quartzite did not accentuate the debitage patterns noted for other materials like obsidian; in fact, it showed a very unique, generalized pattern that was virtually the same for all operations, regardless of which of the two basic technological strategies was employed.

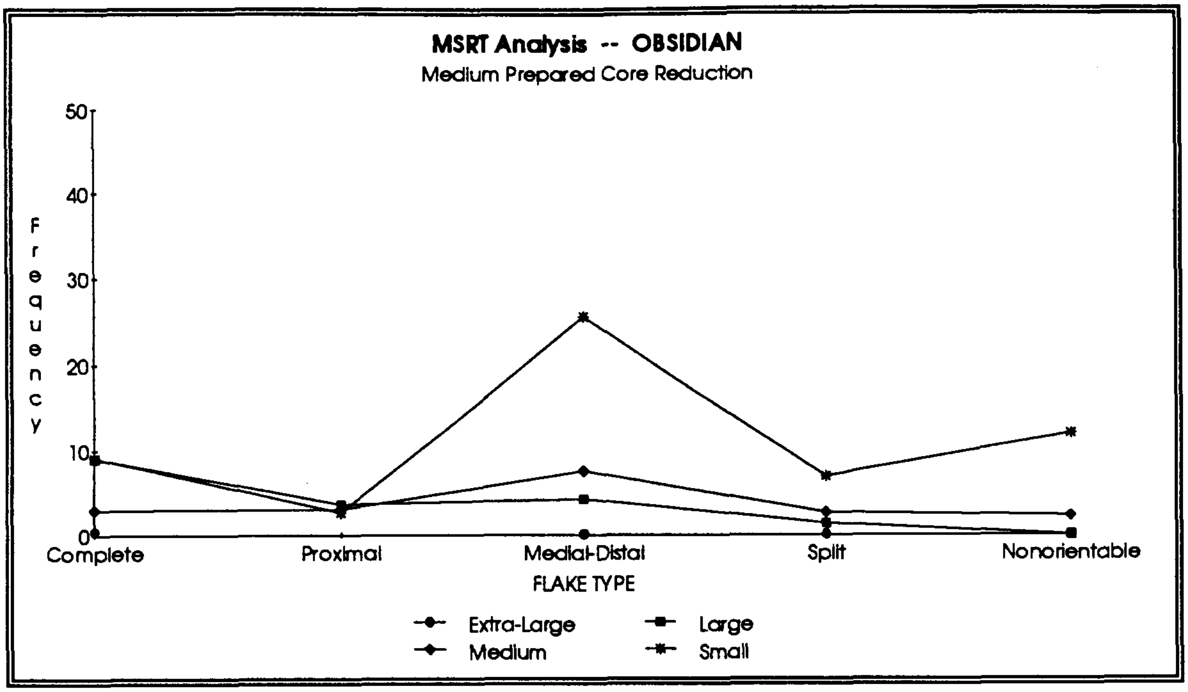
a) Does Core Reduction Produce a Distinctive Debitage Profile?

In general, I found that there was such a degree of variability between debitage from the two sizes of prepared cores and also between the two lithic material types that no single trend in the data could reliably be used to differentiate between core reduction and biface manufacture.

Medium-sized prepared block core reduction produced the closest thing to a distinctive assemblage, since it showed comparable percentages of flakes, by size and type, and also for both raw materials (see Figure 6). This kind of reduction shows the most similarities between the two lithic types. The profile for medium core reduction consisted of elevated counts of small medial-distal and nonorientable fragments, while all other categories showed no patterning. Quartzite did have twice as many proximal fragments as obsidian, however, as well as 5% more "debris", or nonorientable fragments. Obsidian also produced slightly more of the medial-distal pieces in both the small and medium flake sizes. Overall, however, these trends do not match those proposed by Prentiss and Romanski (1989:92), who predicted moderate numbers of complete flakes and medial-distal fragments and high percentages of nonorientable debris.

Compared to medium-sized core reductions, large prepared block core reduction showed drastic differences between the two raw materials (see Figure 7). The quartzite assemblage is composed mostly of nonorientable fragments (33%), most of which are small. The next most common flake class is the small medial-distal one. All other flake categories show no patterning, as was the case for almost all of the reduction operations. Obsidian debitage was dominated by small medial-distal and complete flakes, together comprising 56% of the entire assemblage, on average.

With respect to Prentiss' (1993:130-131; Prentiss and Romanski 1989:92) predictions for core reduction assemblages, extra large and large complete flakes are in fact most common in the large prepared core reduction operations, while the medium-



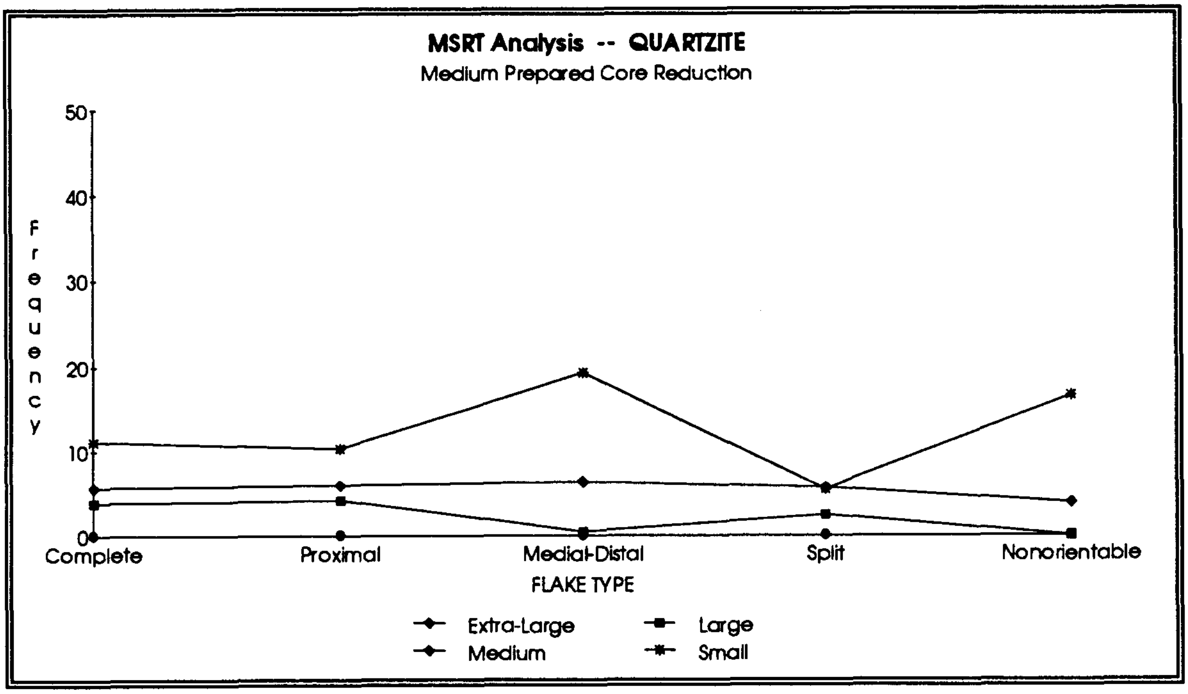
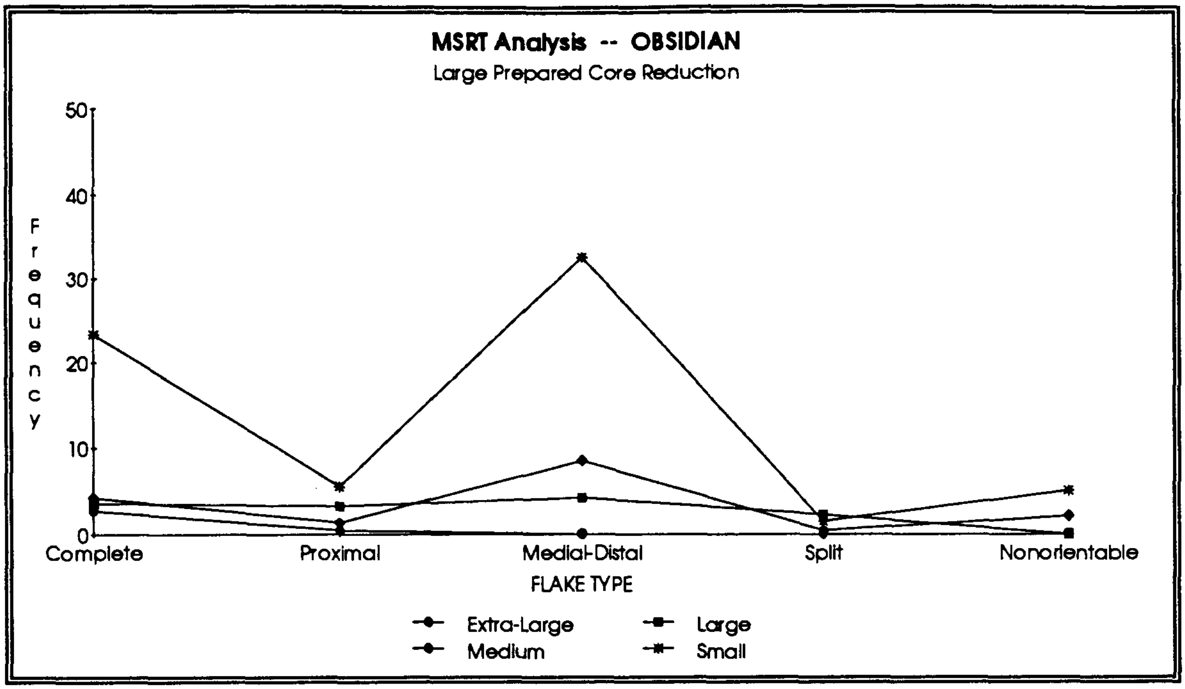


Figure 6: Distributions (by flake type and by flake size) of experimental debitage from obsidian and quartzite Medium-Sized Prepared Core Reduction, according to the Item Completeness (MSRT) Approach.



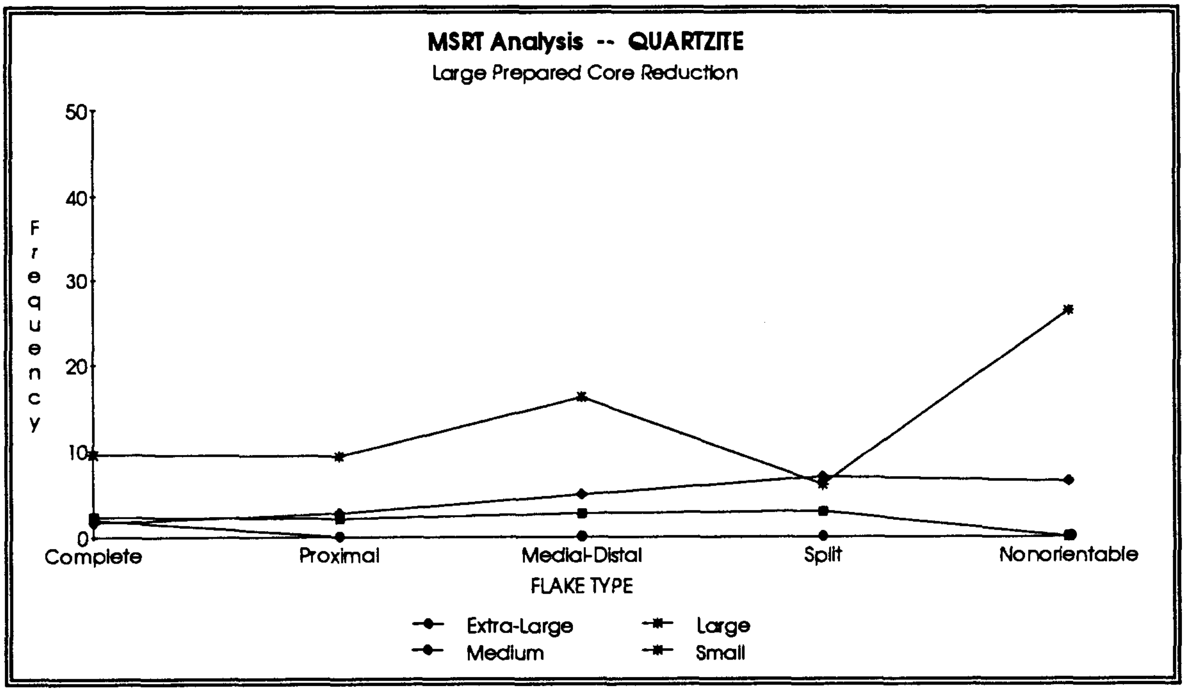


Figure 7: Distributions (by flake type and by flake size) of experimental debitage from obsidian and quartzite Large Prepared Core Reduction, according to the Item Completeness (MSRT) approach.

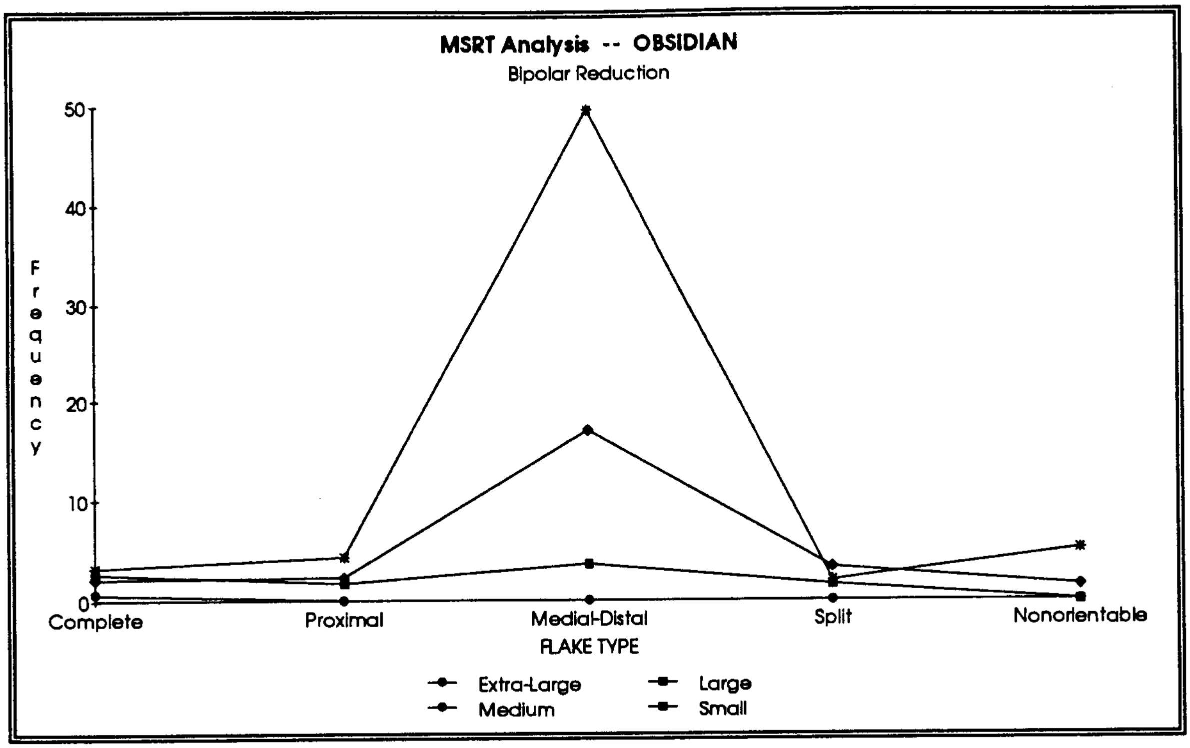
sized cores contribute only small amounts of large flakes. These assemblages were dominated by medium and small flake classes which are thought to be present in all types of reduction activities. In the quartzite assemblages, there are some similarities to the breakage class predictions, whereas the obsidian does not have the "moderate amounts" of complete flakes and medial-distal fragments that were expected, nor does it have great numbers of nonorientable debris (it only had a total of 7% debris). As with the quartzite, the other flake classes show no patterning and are all present in low numbers, at less than 10% for each one.

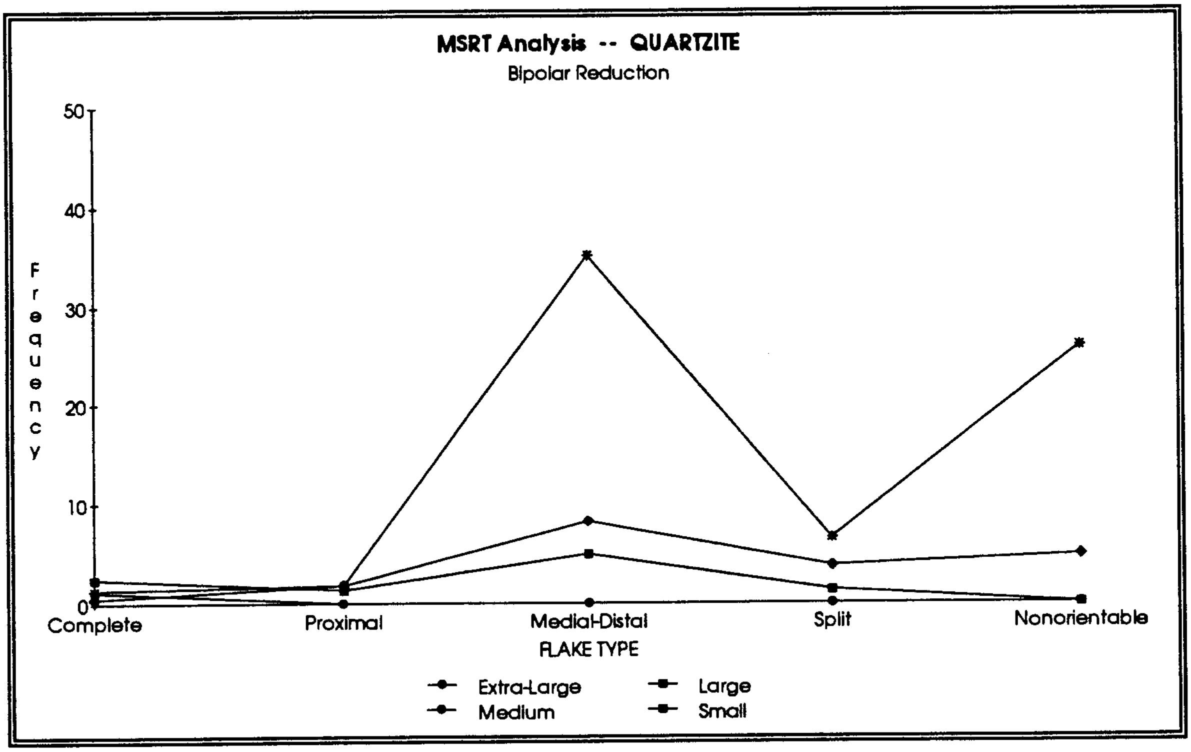
b) Does Bipolar Reduction produce a Distinctive Debitage Pattern?

Of all my reduction experiments, the bipolar ones produce the most distinctive debitage patterning. In this sense, there are obvious indicators that clearly separate bipolar core reduction from biface reduction flaking debris, such as the extremely high proportions of small medial-distal fragments (which make up 35% of the quartzite debitage and 50% of the obsidian debitage). Furthermore, these assemblages (see Figure 8) are also dominated by small nonorientable fragments, followed by numerous medium-sized medial-distal fragments, as noted by Prentiss (1993:431-432) in his vitreous trachydacite ("basalt") experiments.

In addition to these trends, I also observed that the complete flake count was markedly lowered in bipolar reduction (especially the small and medium categories), a pattern that is unique in the obsidian experiments. With regard to quartzite, bipolar reduction was the only case where the small medial-distal fragments are so dominant.

While these debitage patterns are distinct, compared to Stage III biface thinning, the differences between the two raw material types is quite noticeable. For example, even though both lithic types show elevated small medial-distal counts, the obsidian value is





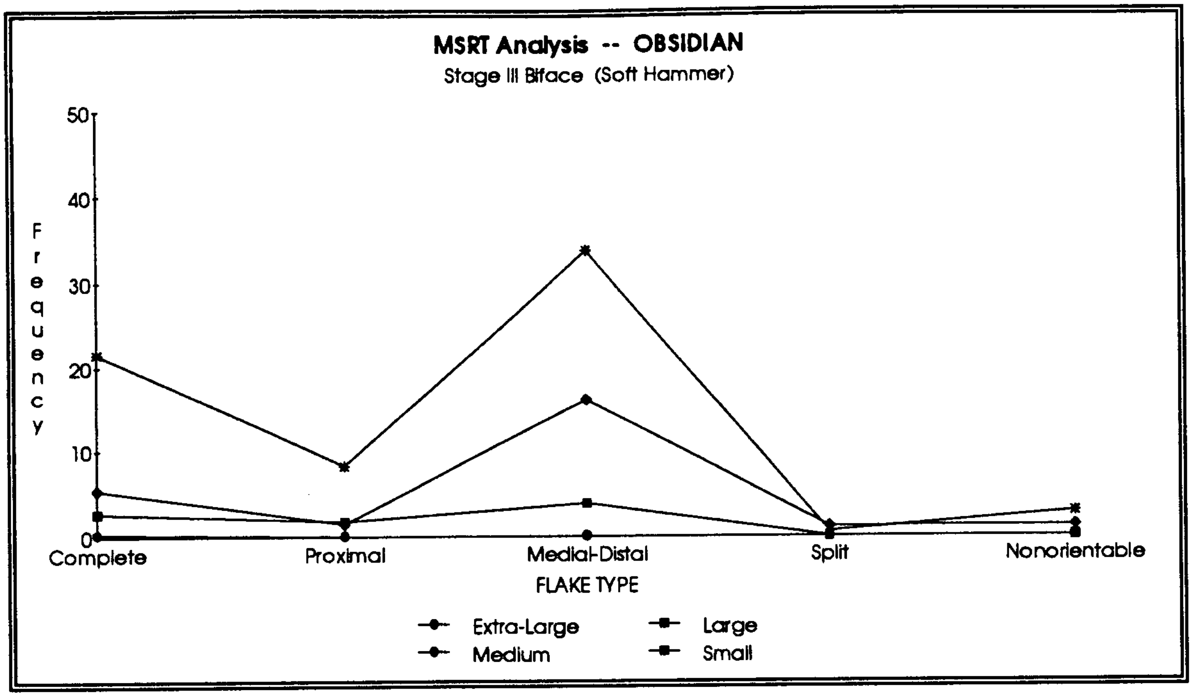
<u>Figure 8</u>: Distributions (by flake type and by flake size) of experimental debitage from obsidian and quartzite Bipolar Reduction, according to the Item Completeness (MSRT) approach.

still 15% higher than quartzite, and is, in fact, the highest count of any flake type in all the experiments. In addition, the percentage of small nonorientable fragments is markedly divergent. With quartzite, this flake type is 21% higher than with obsidian, and is the highest count of this type in all the assemblages. Therefore, while there are distinct patterns for this type of reduction, in comparison to biface thinning debitage, there are still large differences between the obsidian and quartzite flaking debris which make each assemblage unique. Consequently, it is not possible to develop a general debitage pattern or profile for bipolar core reduction.

c) Does Stage III Biface Thinning have a Distinctive Debitage Profile?

Once again, the differences between quartzite and obsidian are quite drastic, such that it is not possible to develop a general debitage profile for tool production in terms of flake breakage categories.

Interestingly, with obsidian, the average flake class distribution for soft hammer percussion (see Figure 9) is virtually the same as that for large prepared core reduction. It is dominated by almost identical quantities of small medial-distal fragments and complete flakes, and less than 5% each for nonorientable flakes and split flakes. The only difference is in the slightly elevated amounts of medium-sized medial-distal flakes, a trait that seems to appear in late stage biface reduction activities. This is seen in the quartzite Stage III biface thinning operations, as well, although the rest of the flake classes are very different from the average obsidian Stage III assemblage. Once again, the quartzite flaking debris is dominated by a generalized pattern of numerous small medial-distal and nonorientable fragments, while the other categories account only for roughly 10% each. On the contrary, Prentiss (1993:153) suggested that nonorientable fragments would occur very rarely in soft hammer biface assemblages. Another noticeable difference between the obsidian and quartzite billet flaking debitage is that split flakes are more commonly found



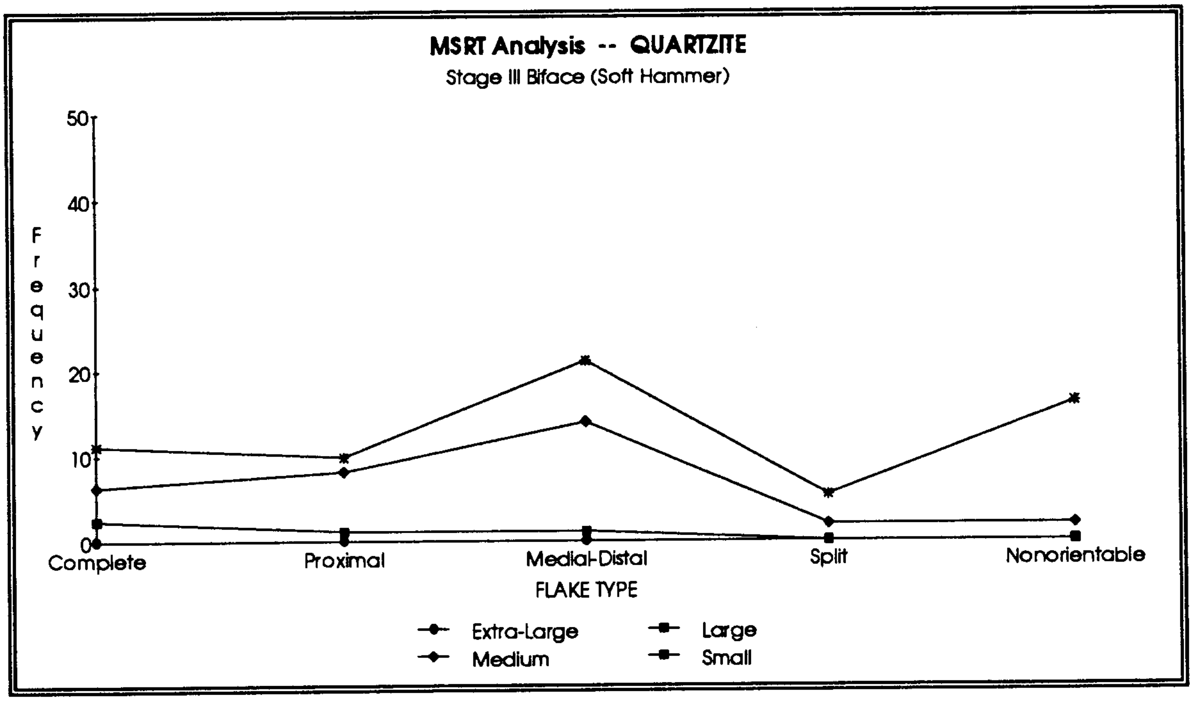


Figure 9: Distributions (by flake type and by flake size) of experimental debitage from obsidian and quartzite Stage III Biface Thinning (soft percussor), according to the Item Completeness (MSRT) approach.

in quartzite than in obsidian assemblages, averaging 7% and 2% respectively.

The flake breakage profiles for this kind of operation are probably the most divergent patterns between the two material types. Furthermore, they do not match the expectations which called for high numbers of complete flakes and equal numbers of proximal and split flakes. Instead, I found that quartzite had almost equal numbers of complete and proximal flakes (20% and 19%, respectively) and an average split flake count of only 7%. In comparison, the obsidian biface assemblages contained an average of 30% complete flakes, 11% proximal flakes, and only 2% split flakes. This profile does not match that predicted by Prentiss and Romanski (1989:92), nor does the quartzite profile.

In terms of flake size, both the extra large and the large categories are extremely rare in tool production, whereas medium and, particularly, small flake sizes are the most common and the most influential as far as patterning goes. In this respect, Prentiss's (1993:130-132) predictions are met. Unfortunately, the overall patterning across these flake sizes does not provide accurate characterizations of the reduction technology employed.

The relationship of core and biface debitage to raw material factors

The hypothesis that each of the two reduction technologies produces distinctive debitage patterns was negated because I found little evidence to support such a claim. Based on this experimental work, it would appear that the overlap between flake profiles for the different reduction strategies, together with the vast differences between raw material types, would not allow accurate or valid interpretations of debitage assemblages.

Raw material characteristics are largely responsible for these discrepancies. Other researchers have suspected that this factor would affect the flake breakage frequencies, along with a series of other influences such as initial core form, idiosyncratic differences between knappers, and skill level (eg. Mauldin and Amick 1989:84; Ingbar et al 1989:121;

Prentiss and Romanski 1989:173); however, little testing has been carried out to assess these important variables.

Since this approach is based on the completeness of flaking debris, it is obvious that the physical and petrological attributes of a lithic material will determine how it fractures. This fracture behaviour, in turn, determines the morphology of the debitage and therefore its classification according to the Sullivan and Rozen or MSRT approach. For this reason, it is vital to take into consideration the physical properties of each material being examined, since these traits will be transposed directly into the flake-class frequencies.

One of the most influential physical characteristics of a lithic material, in relation to debitage analysis, has proven to be internal flaws and inhomogeneities such as inclusions or incipient fracture planes. With regard to the item completeness approaches, this factor can reduce the number of flakes that are categorized as "complete", since this flake type is based on whether or not the flake margins are intact; this, in turn, is determined by the type of distal termination on the flake. If the flakes have either a feather or hinge termination, they are clearly considered to be complete; if lateral snaps or breaks do not interfere with accurate width measurements, then the flake is also considered "complete" (Sullivan and Rozen 1985:759). However, if a flake has a step termination, it would not be included as a complete flake, but rather as a proximal or "broken" fragment.

According to Cotterell and Kamminga (1987:700), step terminations are often caused when the crack encounters a flaw in the material which serves to blunt the propagation phase and cause the crack to change directions; this is referred to "crack arrest", and they note that this is a common occurrence in "highly flawed materials such as quartzite". Luedtke (1992:86) also noted the production of broken flakes caused by inhomogeneities. In this respect, flaw-induced step terminations in my experimental quartzite cores and bifaces may be largely responsible for increasing the numbers of proximal flakes and perhaps medial-distal fragments, as compared to obsidian. Figures 6,7,and 9 show how proximal flakes and complete flakes occur in almost equal amounts in quartzite core and biface operations, whereas, with obsidian (which is generally quite homogeneous), there are two or usually three times more complete flakes than proximal ones, on average, since the fractures are not stopped by internal flaws.

So-called "poor quality material" like quartzite, with its incipient fracture planes, voids, and inclusions, might also be expected to increase the number of nonorientable fragments (Baumler and Downum 1989:107). This debris is attributed to the shattering of striking platforms and bulbs of percussion, and is expected to become more abundant if core reduction becomes more intense (Sullivan and Rozen 1985:763). Since heavier hammerstones and more force were required to remove flakes from quartzite prepared cores, it makes sense that this is reflected in the higher amounts of nonorientable debris found in quartzite debitage assemblages. This phenomenon explains the differences between my quartzite and obsidian core reduction debitage profiles which comprise, on average, 20% nonorientable fragments for quartzite medium core reduction versus 14% for obsidian, and 33% nonorientable fragments in quartzite *large* core operations, compared to 7% for the same activities employing obsidian. Furthermore, in bipolar reduction, the nonorientable fragments make up 31% of the quartzite debitage but only 7% of the obsidian assemblage. As mentioned in the discussion of Hypothesis # 2, the use of more intense force and heavier percussors also explains the greater frequency of split flakes in the quartzite collections, particularly in all the hard hammer operations.

To summarize, then, the greater frequency of internal flaws and the overall inhomogeneity of quartzite is probably responsible for the different flake-class frequencies based on flake breakage or item completeness. As a result, quartzite differs significantly in its physical characteristics from other homogeneous materials like obsidian and, consequently, it is difficult to develop an average or standard pattern that differentiates between biface and core reduction debitage assemblages. In addition, there is a fair amount of variation, even within one lithic material type, caused by different core morphologies or other factors. In general, as Prentiss and Romanski (1989:97) also suspected, biface or core reduction assemblages may be more variable than expected, therefore the effects of factors such as raw material properties must be explored experimentally, not assumed.

EVALUATION OF HYPOTHESIS #4

Hypothesis 4: The flake aggregates produced by different technologies and in different stages of manufacture will exhibit markedly different size grade distributions for both obsidian and quartzite (after Ahler 1989b:205).

Expectations

Ahler's experimental research has been underway for over twenty years and has been subjected to numerous statistical analyses, all of which have consistently indicated a high degree of reliability unmatched by other types of procedures or classification schemes. Other analysts have also put this method to the test in their own experiments and have found it to be particularly successful in isolating stage variation in biface reduction debris (Odell 1989:184; Mauldin and Amick 1989:73).

Nevertheless, I was uncertain about how successful it would be in interpreting quartzite debitage assemblages, since there has been virtually no mention of flake aggregate analysis experiments using this kind of material. Stahle and Dunn (1982:94) have suggested that the potential effects of factors such as raw material differences on the flake size distributions should be examined through controlled replicative experiments but, up until the present, the primary basis for Mass Analysis comparisons has relied on cherts, especially Crescent chert and Knife River flint, for the bulk of their inferences. At the outset of my experiments I presumed that the general trends for the chert debitage would be close enough to allow comparisons to the Morrison quartzite.

In general, an analyst using this method is told to expect that late stage biface thinning operations are distinct from earlier edging or preliminary reduction in terms of differences between the flake aggregates in the four size grades. Furthermore, core reduction debitage is expected to differ noticeably from bipolar or other flaking debris with regard to flake class distributions. While Ahler does not provide many specific predictions for how these trends will materialize in counts, weights, or ratios across the size grades, he does provide generalized data from his experimental project for comparison. One must expect, therefore, that his comparative profiles cover the full range of possible variation and that debitage from other materials will be close enough to enable reasonably accurate extrapolations.

The assumptions upon which mass analysis is based include the ideas that knapping is a reductive process, and that variation in load application affects flake shape (Ahler 1989b:93). For example, debitage derived from nonmarginal flaking (associated with freehand core reduction and edging operations) can be expected to have greater thickness and length, therefore it will likely have a greater average weight in comparison to marginal, or billet flakes (Ahler 1989b:91). Since flakes within a given size grade will be of the same general morphology, the average *weight* of flake aggregates is an important measure because it may allow the differentiation of marginal from nonmarginal debitage. Bipolar flakes should also have a greater mean flake weight than that of bifacial thinning debris. These differences should be evident even if the flake counts are the same for the size grades, although flake counts and ratios of flake counts by size grade are also thought to exhibit distinct profiles depending on the type of technology used in the operations and the stage in the manufacturing sequence.

With regard to cortex cover on debitage, the relative frequency of cortical flakes can be expected to be highest in early and single-step manufacturing operations, and lowest in late stage and complex procedures (Ahler 1989a:90). However, individual cortex profiles can vary greatly from one lithic material to the next, since some types occur in fully corticate cobbles or nodules, while quarried material will generally be free of cortex cover. Cortex profiles can also vary according to the initial size of the material, the knapper's goals, the reduction technology employed, and the stage of manufacture, each of which should create a unique distribution of the debitage across the size grades (Ahler 1989a:90).

Results of My Experiments

Hypothesis 4: Partially Supported/Partially Negated

The flake aggregates produced by different technologies (such as hard hammer, soft hammer, or bipolar reduction) and at different stages of manufacture (as in early, middle, and late bifacial reduction) showed marked differences only in some of the variables recorded for the four size grades. Variables related to weight, such as the percentage of the total debitage weight in each size grade or the mean weight of flakes in each grade, accurately differentiated the various reduction strategies and stages; however, variables dealing with flake counts (eg. percentage of flakes in each size grade or the ratio of G4:G1 to G3 counts) were not successful, overall. Finally, relative cortex frequency in each size grade was so unpredictable that this variable only showed marked differences between the reduction operations when averages of the reduction operations were examined.

In general, I found that the differences between the obsidian and the quartzite debitage profiles are quite extreme; therefore, it would appear that direct comparisons from one type of raw material to another will not produce accurate or valid inferences on the origins of an assemblage. In this sense, the use of mass analysis data for comparative purposes should be carefully considered prior to applying it to an untested lithic material type. Raw material differences also affected the success of the approach in several ways, although this will be discussed later in more detail.

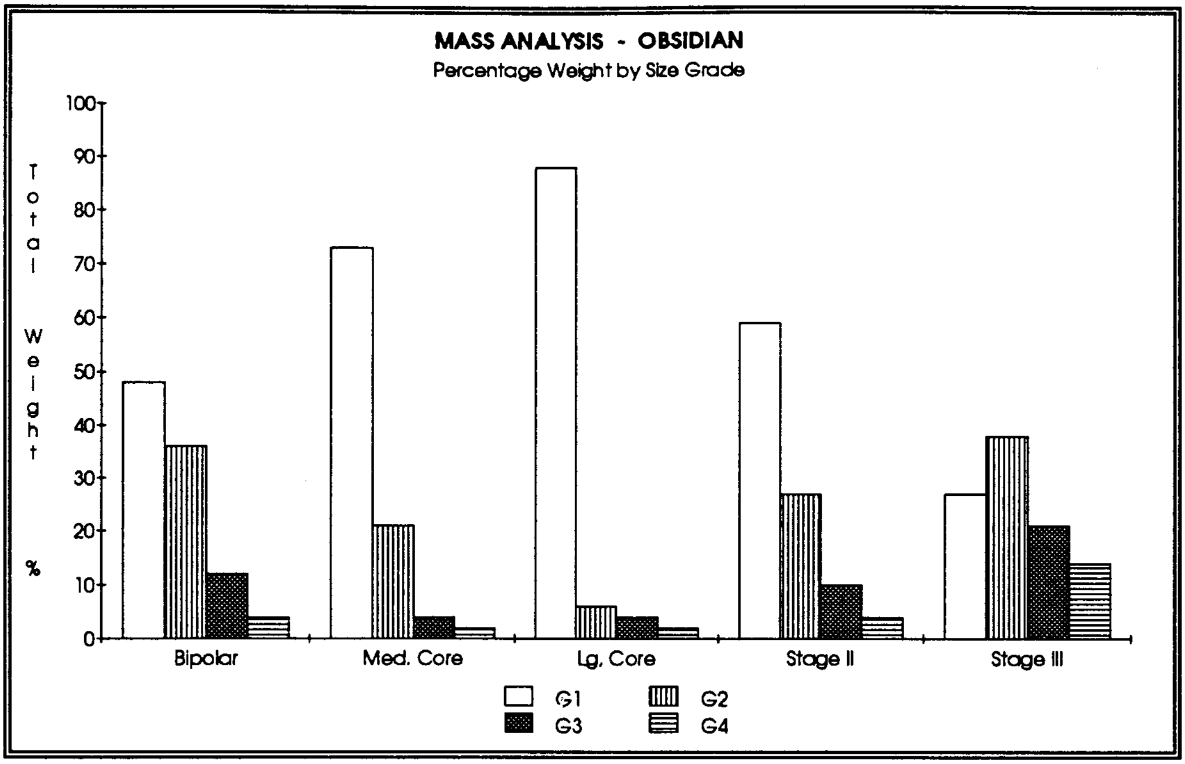
The other aspect of my assessment of this approach was to determine the success or failure of several specific variables used in connection to the size grade distributions. These will be examined individually, below.

a)Weight Variables

Percentages across size grades according to the weight of the debitage classes (see Figure 10) proved to be a useful variable for distinguishing between the different reduction technologies and stages of biface manufacture employed in my experiments.

The quartzite and obsidian Stage III biface thinning assemblages show a very unique flake-weight profile that differs from the others by being the only case in which most of the weight was in the G2 size class, rather than in G1. Furthermore, the debitage weight is more evenly distributed among the smaller size grades, since the G3 and G4 classes are noticeably elevated in comparison to the other operations (particularly core reduction). In this sense, the mass analysis approach is a success, since it clearly separates the flaking debris of late stage billet thinning from that of hard hammer edging, prepared core reduction, and bipolar reduction for both obsidian and quartzite.

However, as with most of the other variables, this success is quite limited and the results can often be ambiguous. For example, while Stage II biface edging debitage (from hard hammer percussion) is also distinct from the others in certain ways, this difference is not very marked, especially in the quartzite assemblages, where all of the hard hammer operations, including bipolar reduction, are extremely similar to one another. Therefore this weight variable can separate late stage soft percussor debris from hard hammer



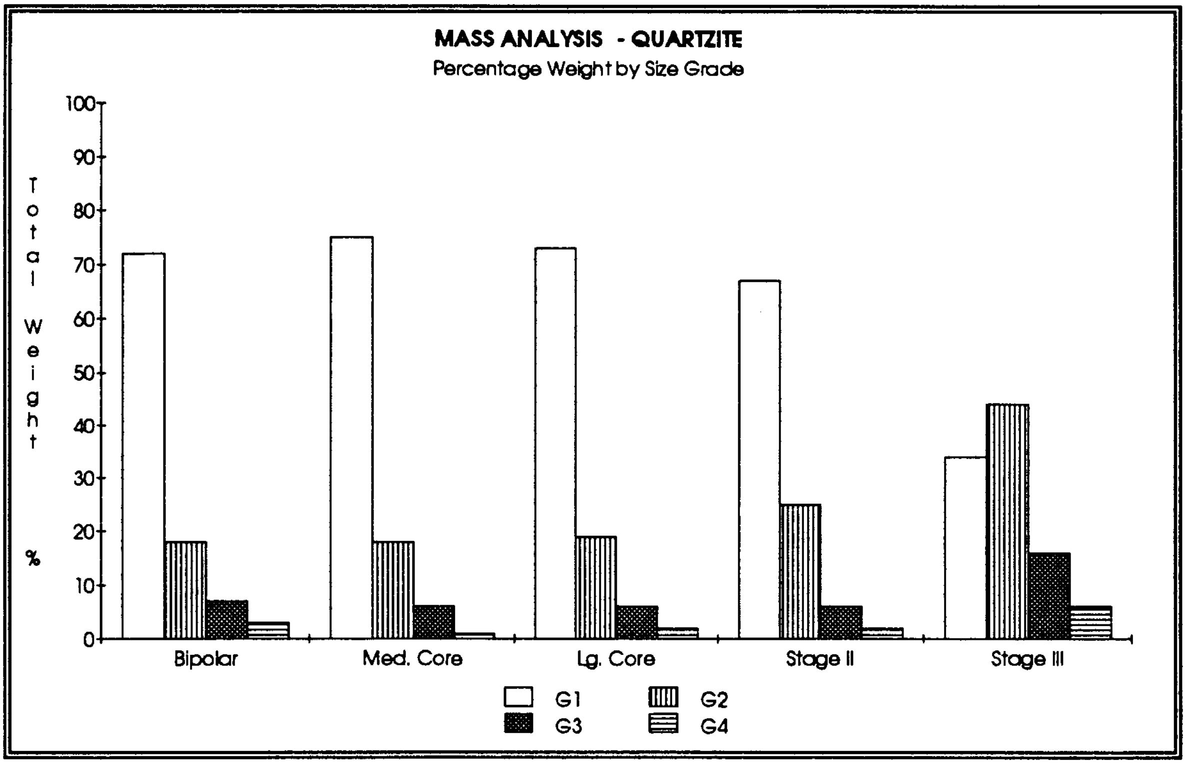


Figure 10: The percentages of flakes by weight across the size grades for obsidian and quartzite experimental debitage, according to the Mass Analysis approach.

technologies, but it could not isolate the middle stage of biface manufacture with a great deal of success.

The other weight variable examined is the mean flake weight in each size grade (see Table 6). This variable provides the most reliable results of all the mass analysis variables and is able to show distinctive patterning for all the different kinds of reduction technologies, including bipolar and block core reduction, as well as between the middle and late stages of biface manufacture.

With few exceptions, the prepared block core operations produce the highest average flake weights across the three largest size grades, followed by bipolar core reduction, then Stage II biface edging, and finally Stage III soft hammer thinning. The G4 (smallest) size grades, however, show almost no distinctions whatsoever, with the exception of the quartzite bipolar assemblages. This general pattern is consistent in both quartzite and obsidian assemblages, although between the two raw material types there are still significant differences in the actual mean flake weights, with quartzite flakes always being much heavier than obsidian ones.

Overall, the biggest differences between technologies and reduction stages are in the larger (G1 and G2) classes, with the distinctions becoming less and less pronounced as the size grades get smaller. Stage III billet flakes tend to weigh less than half of the middle stage edging debitage, which is markedly lighter than the average block core flakes. In some cases, the core reduction mean flake weight is five or six times that of the late stage billet flakes. This type of patterning is quite robust, which I found to be unique among the variables studied in this approach.

MEAN FLAKE WEIGHT (g) BY SIZE GRADE -- OBSIDIAN

	G1	G2	G3	G4
Bipolar	33.2	4.4	0.57	0.06
Bipolar	28.7	5.8	0.43	0.04
Med. Core	32.4	6.6	0.4	0.05
Med. Core	42.8	4.9	0.27	0.02
Large Core	57.2	4.4	0.37	0.04
Large Core	68.8	3.5	0.43	0.08
Stage 2	23.1	2.7	0.52	0.04
Stage 2	30.9	3.9	0.4	0.05
Stage 3	12.5	2.6	0.31	0.05
Stage 3	10.2	2	0.27	0.04
Averages:				
Bipolar	31	5.1	0.5	0.05
Med. Core	37.6	5.8	0.34	0.04
Large Core	63	4	0.4	0.06
Stage 2	27	3.3	0.46	0.05
Stage 3	11.4	2.3	0.29	0.05

MEAN FLAKE WEIGHT (g) BY SIZE GRADE -- QUARTZITE

	G1	G2	G3	G4
Bipolar	65.6	5.6	0.57	0.07
Bipolar	67	3.4	0.58	0.12
Med. Core	105.4	4.3	0.55	0.05
Med. Core	99.6	6.1	0.71	0.07
Large Core	68.2	5.8	0.57	0.07
Large Core	83.1	5.7	0.53	0.06
Stage 2	40.6	2.3	1.19	0.06
Stage 2	16	3.8	0.57	0.06
Stage 3	24.9	4.2	0.5	0.06
Stage 3	14.1	2.8	0.51	0.06
Averages:				
Bipolar	66.3	4.5	0.58	0.1
Med. Core	102.5	5.2	0.63	0.06
Large Core	75.7	5.8	0.55	0.07
Stage 2	43.3	3.1	0.88	0.06
Stage 3	19.5	3.5	0.51	0.06

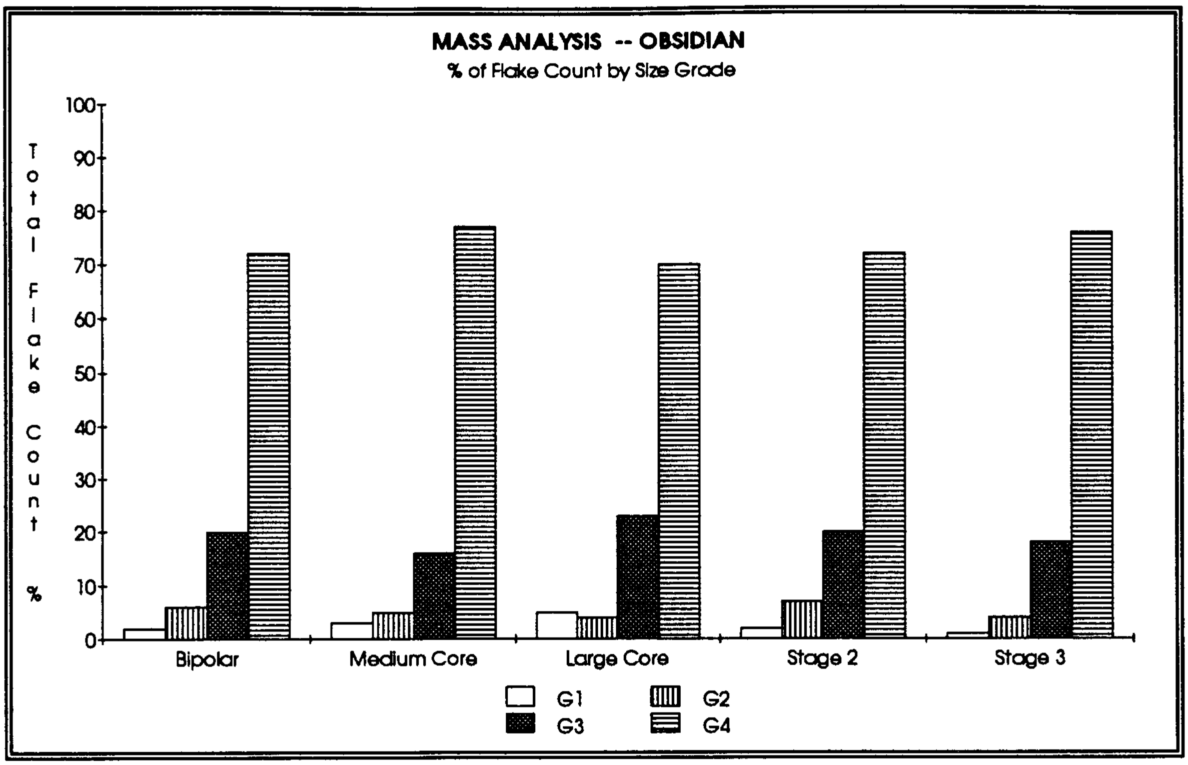
TABLE 6: Mean Flake Weights (in grams) by Size Grade for obsidian and quartzite experimental debitage, according to Ahler's Mass Analysis approach.

b) Flake Count Variables

Unlike the weight variables, which seem quite reliable and accurate, the two variables that dealt with flake counts achieve little or no success in differentiating between the various reduction strategies or stages.

The percentage of flakes by count in each size grade showed virtually identical profiles for all operations, whether they were billet flaking or bipolar core reduction (see Figure 11). Quartzite, however, shows more variation across the size grades, although there are not enough differences between the technologies or manufacturing stages to allow their distinction. All assemblages are dominated by the flakes in the smallest size grade (G4), followed by moderately high percentages of G3 flakes. The largest size classes are relatively uncommon compared to the small debitage, although there are more of them in quartzite than in obsidian samples. Essentially, this variable is not successful at all.

The other flake count variable (the G4:G1 to G3 ratio), which is the ratio of flake counts in the smallest size grade to the combined counts in the other three classes, is intended to capture flake size variation for each of the operations (Ahler 1989b:209) (see Figure 12). However, according to this variable, the only knapping activities that stand out are quartzite bipolar reduction and both medium prepared block core reductions. Aside from these, and the slight elevation of the obsidian Stage III ratio, all other operations have similar ratio values, whether they are from early or late stage biface production or general core reduction.



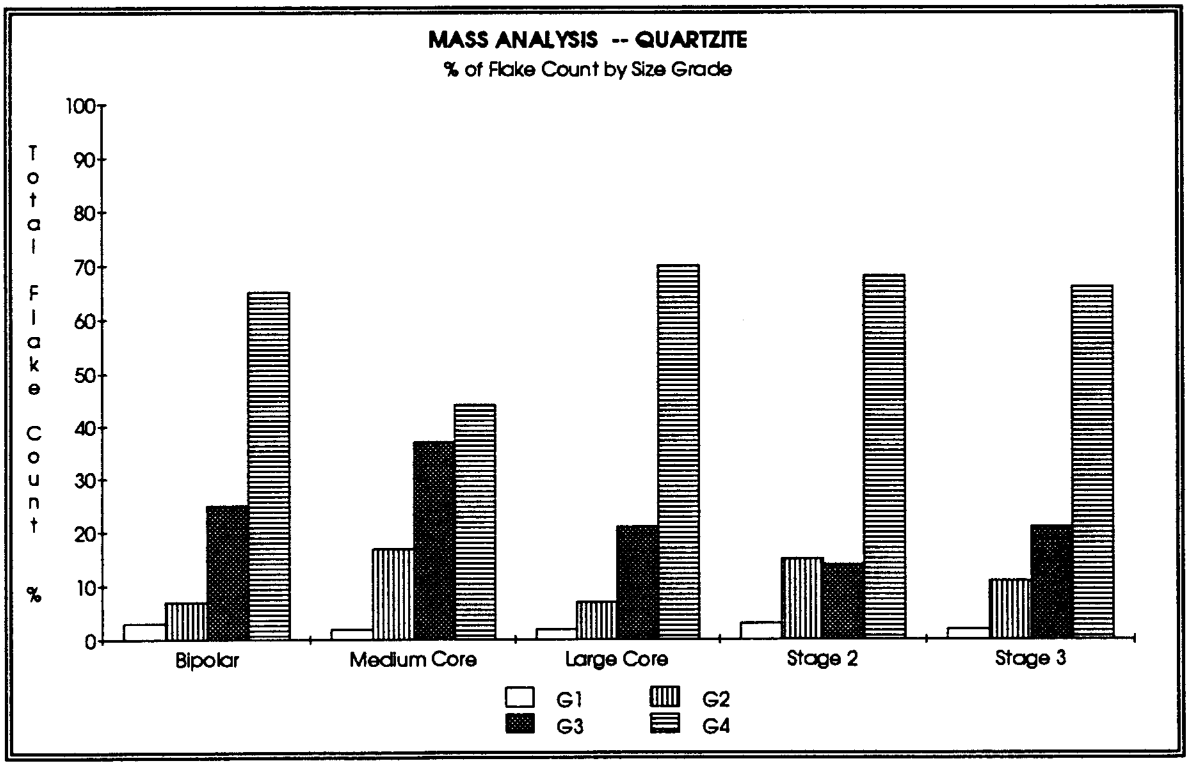


Figure 11: The percentages of flakes by size grade for obsidian and quartzite experimental debitage, according to the Mass Analysis approach.

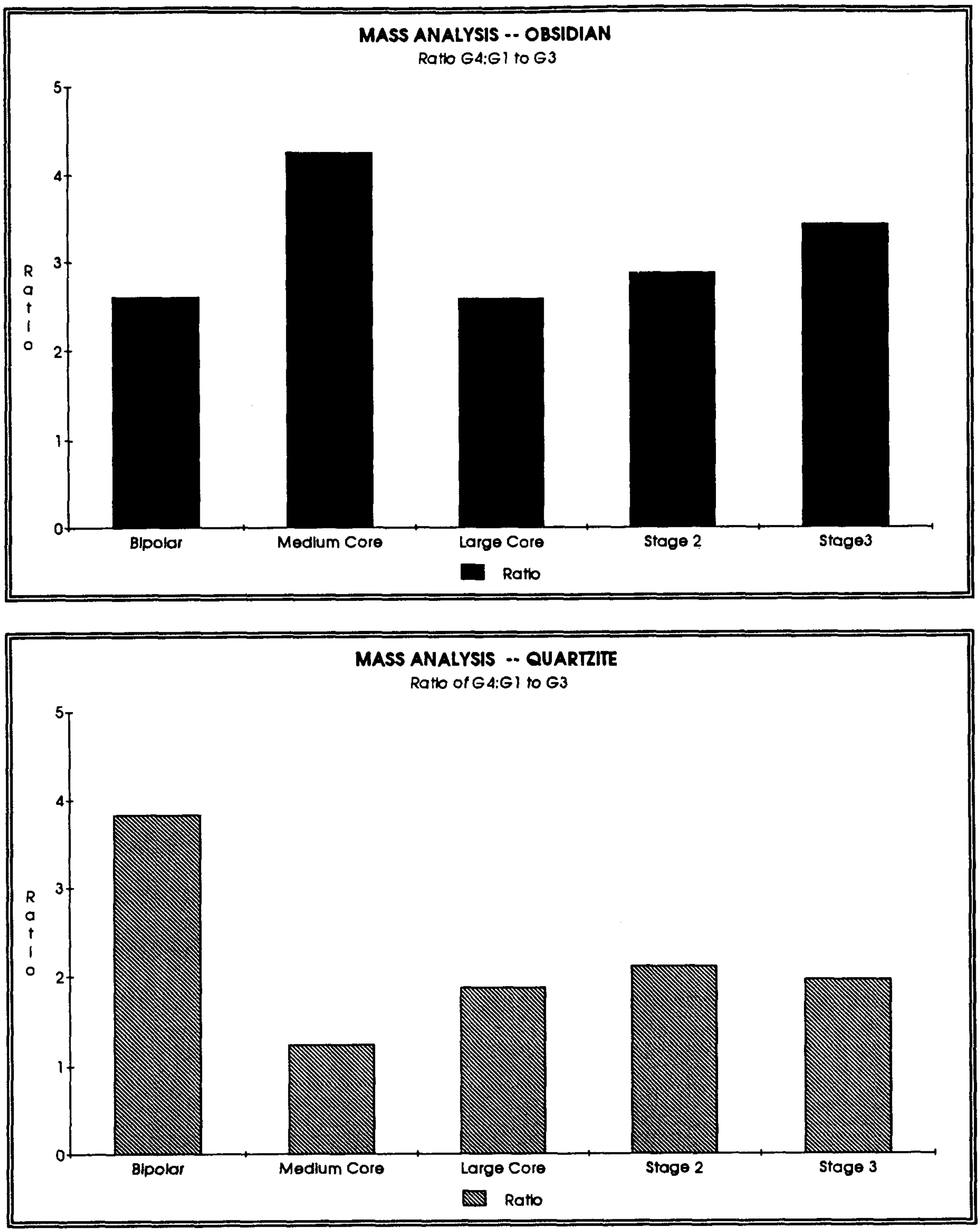


Figure 12: The ratio of G4 flakes to the combined counts of flakes in size grades G1 to G3 for obsidian and quartzite experimental debitage, according to the Mass Analysis approach.

c) Relative cortex frequency

Since cortex is presumed to decrease as reduction progresses, one would expect that late stage bifacial thinning would have the least amount of cortex on its flaking debris. This phenomenon is just what I found in the average percentages of cortical flakes by size grade, for both quartzite and obsidian (see Figure 13). Stage III obsidian flakes have extremely low frequencies of cortex in the three smallest classes and even the G1 class is the lowest of all the various knapping activities. Stage II edging operations have unique average percentages of G2 and G3 cortical flakes, although there is not a clear distinction of this reduction stage from the other types of reduction events like there is in the quartzite samples. With quartzite, both of the biface reduction and thinning operations show distinct average profiles of cortical flakes in comparison to the other hard hammer core reduction technologies which all showed similar and standard types of profiles.

In this case, the mass analysis cortex variable is successful in distinguishing the technological and manufacturing origins of the flaking debris. The greater degree of success occurs with the quartzite, since all five of the cortex profiles for the obsidian operations are different (even each of the hard hammer reduction events which involve the same knapping techniques). Once again, there is some success at differentiating certain activities from others, but this is not as reliable for obsidian collections. The differences between the two raw material types is again quite extreme, probably due to the different initial forms of the materials, and calls the validity of this variable into question.

The relationship of the Mass Analysis results to raw material factors

Even though the *averages* for the relative cortical frequencies show that there are marked differences between the knapping events, an examination of the full range of reduction operations illustrates just how much variability there really is (see Figure 14). In fact, the variability is so great that, only when the averages are considered, are there any

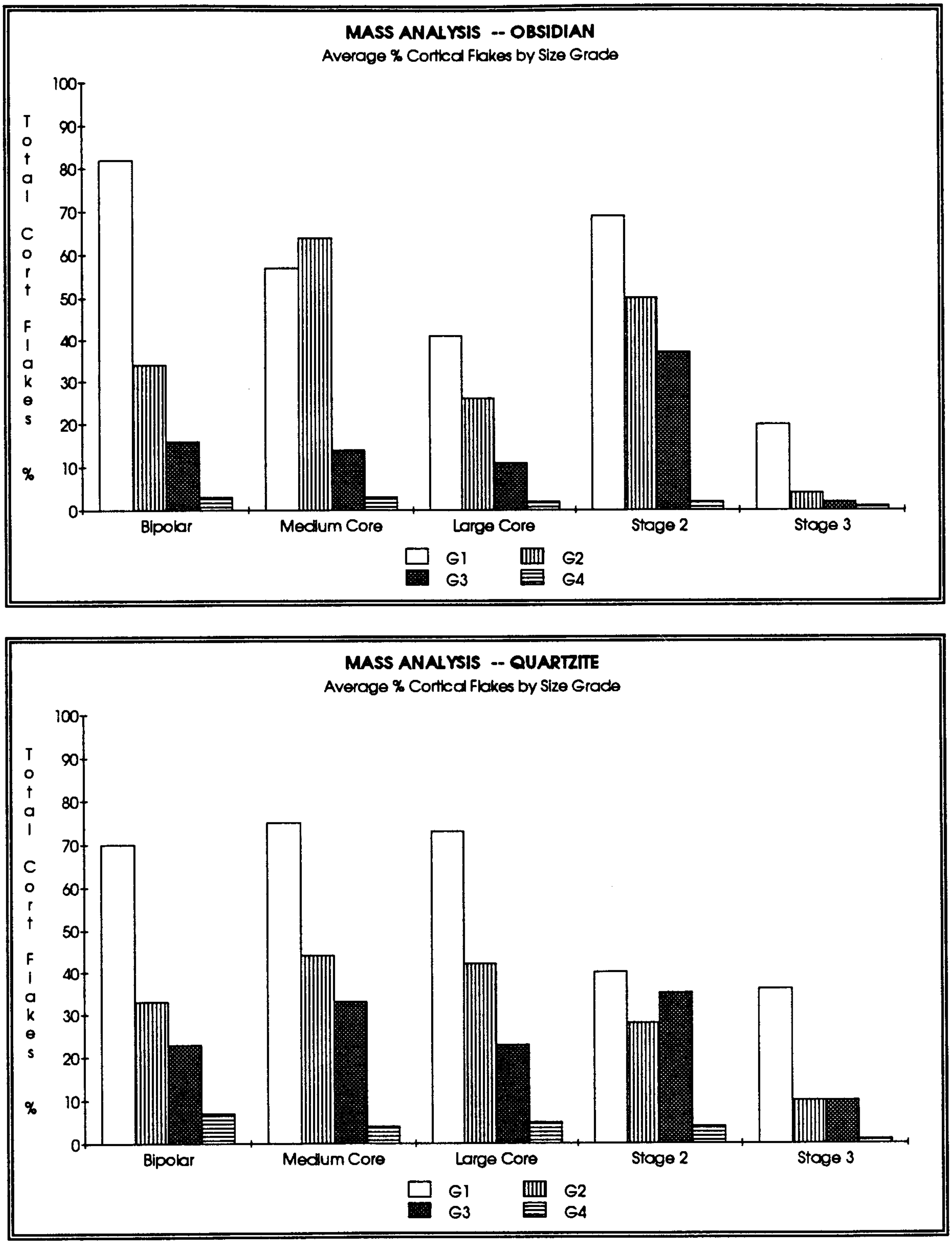
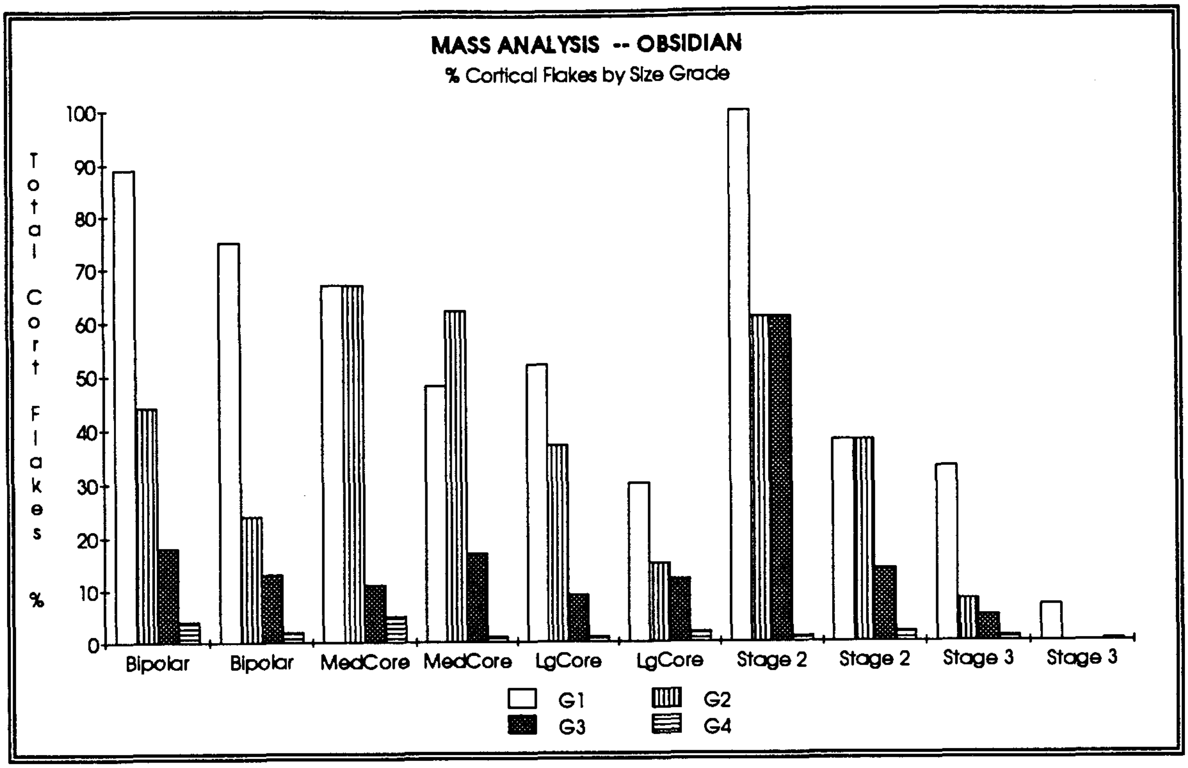


Figure 13: The average relative frequency of cortical flakes by size grade for obsidian and quartzite experimental debitage, according to the Mass Analysis approach.



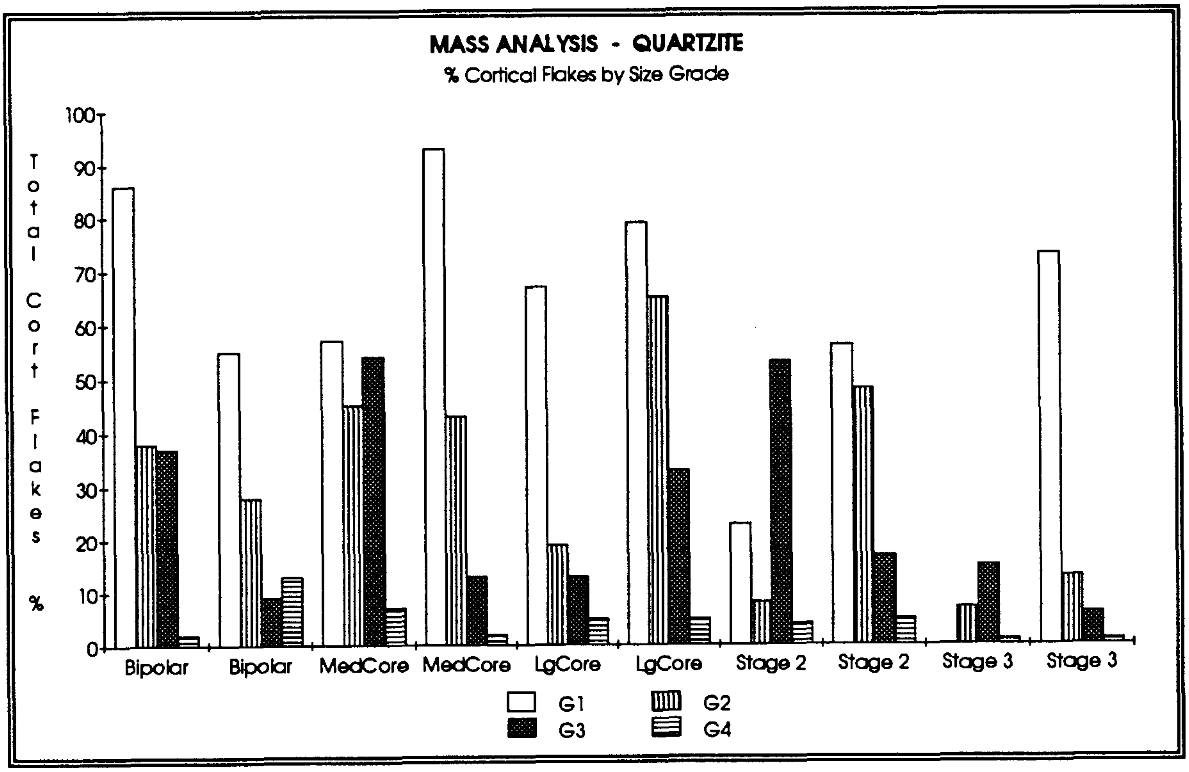


Figure 14: The relative frequency of cortical flakes by size grade for all obsidian and quartzite experimental debitage, according to the Mass Analysis approach.

visible trends in the size distributions of cortical flakes. This phenomenon demonstrates one important factor that is often overlooked in debitage analysis: that the original size and morphology of the material will determine the amount of cortical flakes in the assemblage. In my experiments, both the obsidian and the quartzite are from highly different sources and are covered in highly variable amounts of cortex. For example, the obsidian from Glass Buttes, Oregon, is generally procured in nodule form in a range of sizes, although the amount of cortex is quite variable on these cores. Sometimes it is a hard, thick rind, while other times it is merely the dulled or smoothed-over surface of a previous fracture. In addition, the quartzite from Wyoming is also quite variable since it is procured from surface exposures of varying size and degree of protuberance; consequently, the cortex cover, which is essentially a thin, weathered layer, can be vastly different or absent altogether, depending on where the core or blank was removed from on the exposure, and depending on the blocky or tabular nature of the original piece.

Since these two material types can vary so much in terms of the frequency of cortical flakes, it is no wonder that this variable showed such a wide range of profiles for the reduction activities. These reasons might explain why it sometimes succeeds and other times fails to reliably differentiate between the technologies and stages of reduction.

The flake count variables also are affected by raw material variability. Some researchers argue that the original morphology of a core or blank can be "the largest factor in determining the characteristics of flake size distribution" (Patterson 1990:554).
All the G4:G1 to G3 ratios of flake count, for example, are generally lower (between 1.3 and 2.0) for quartzite assemblages than for obsidian ones, which ranged from over 2.5 to 3.5. Differences are also observed in the percentages of flake counts by size grade, which look almost exactly the same for the very homogeneous and predictable obsidian, whereas the quartzite profiles for this variable are quite different from one reduction technique to

the next, possibly due to the vast differences in its texture and physical properties, even within a single nodule.

Finally, the differences in the weight variables between the two lithic types are due to differences in their physical properties. In fact, the reason that obsidian shows such drastic differences from quartzite in the percentages of weight by size grade and in the mean flake weight by size class is likely due to the specific gravities of these different materials. Bouey (1983:39) has argued that weight can be misleading because materials with higher specific gravities will weigh more, even though the flakes are the same size. This would help to explain why the profiles from obsidian and quartzite assemblages can look so different even in the same size grades for the same kinds of operations, although it is also possible that the greater toughness of quartzite is responsible for thicker flakes in general.

Since the level of comparability between raw material types is so low, the success of the Mass Analysis approach is limited to the size of its comparative experimental collection and by varying mixtures of stone types in debitage assemblages. Extrapolation of the size grade data from one lithic type to another would therefore likely result in inaccurate interpretations of the collections. So, here too, we see a method of analysis for debitage being negatively affected by factors relating to variation in raw material properties and composition.

SUMMARY

This chapter has discussed the analysis of the experimentally-generated quartzite and obsidian debitage assemblages. Observations on the behaviour of these two materials were included as an adjunct to the evaluation of the four hypotheses that were being tested.

The effect of raw material differences on the success of the three analytical approaches was found to be quite critical. In all cases, the differences between the obsidian and the quartzite samples were great enough to jeopardize the overall validity of each method employed.

Specifically, Magne's Individual Flake Attribute Approach was the most detrimentally affected of all. Not only was the quantification of flake scars found to be an invalid procedure for identifying the technological origins of debitage samples, but so too was the use of so-called technique-specific classes like biface thinning flakes and bipolar flakes based on his particular identification criteria. The influence of petrological factors played an important role in creating the problems associated with this method, since the divergent physical properties of the two materials determined how (or whether) the fracture features manifested themselves in the debitage. As a result, it was found that increased numbers of dorsal scars and named flake types can be produced by a variety of means in large enough quantities to obscure the real reduction technique, and cannot always be correlated with specific or progressive thinning operations. Therefore, the stage of manufacture and the flaking technique employed cannot be inferred from these traits, nor can they be assumed to be free from the influence of raw material factors.

The modified Sullivan and Rozen typology was also found to be affected by differences in the physical nature of the two lithic types. Patterns in the distribution of flaking debris across the flake breakage categories showed few general trends, with the exception of bipolar profiles which were very unique and easily identifiable using this approach. In terms of the original goals of this typology, however, it was not possible to differentiate between prepared core and biface reduction assemblages, particularly when made from quartzite. Physical characteristics concerning flaws and inhomogeneities like incipient fracture planes were more common in the quartzite, and this affected how the flakes were classified according to this approach. Furthermore, factors relating to the different knapping strategies required by the two material types also affected the debitage patterning. In general, the variation both between the materials as well as within each one did not allow for any type of standard or independent patterning that could be used to accurately infer the reduction strategies employed.

Finally, physical differences between obsidian and quartzite had an important impact on Ahler's Mass Analysis approach, although not as severely as in the other two methods. In this case, some of the variables calculated from the flake aggregate data (especially mean flake weight) did indeed show marked differences between the hard hammer, soft hammer, and bipolar debitage, and also between the stages of biface manufacture; on the other hand, flake count variables tended to be unsuccessful. Furthermore, in all knapping operations, the differences between quartzite and obsidian were always quite extreme, therefore it is not possible to draw valid inferences on one raw material type using the data derived from another. In addition, cortex cover also illustrated the divergent nature of the debitage patterns from the two materials, and I found that this variable was too unpredictable to make any accurate generalizations from, even within a single material type.

My findings show that some of the petrological characteristics of a lithic material directly affect the way that it fractures and, consequently, the way that the resultant debitage is patterned, whether this is analysed from the perspective of flake scars on individual flakes, or flake breakage trends in entire assemblages, or from the composition of flake aggregates grouped by size grades. Therefore, the validity of these approaches to debitage analysis is jeopardized, since they were usually unable to measure what they claim to. The full impact of this conclusion will be explored in Chapter Six.

CHAPTER SIX:

IMPLICATIONS OF THE EXPERIMENTAL FINDINGS

This research project was carried out in an attempt to extend the testing of some important assumptions in debitage analysis. In particular, I wanted to determine whether or not the Individual Flake Attribute approach, the modified Sullivan and Rozen typology, and Mass Analysis method could effectively and accurately identify the technological origins of experimental flake assemblages. I wanted to find out if these procedures were internally valid; that is, would they reproduce the same patterns that other researchers had found when applied to a best-case scenario such as an "excellent" raw material like obsidian? Furthermore, would they prove themselves to be externally valid by living up to their expectations when applied *outside* of the original experimental parameters, as in the analysis of debitage from a "poor quality" raw material like quartzite? In essence, I was trying to determine whether or not these three approaches were valid in terms of the theoretical constructs that guide them. Would they actually measure what they were intended to measure, or was it possible that they would produce artificial patterning in the data that would then lead to the development of inaccurate interpretations?

EXPERIMENTAL RESULTS AND THE ASSESSMENT OF VALIDITY

Based on information gained from extensive research into the topics of petrology, fracture mechanics, and replicative lithic analysis, I was able to gain a basic understanding of the physical aspects of stoneworking and debitage formation. This background knowledge enabled me to assess the factors affecting patterning in the various flake classifications used by Magne (1985), Sullivan and Rozen (1985)/Prentiss(1993), and Ahler (1989a, 1989b), when applied to experimental quartzite and obsidian debitage assemblages.

Internal Validity

Since my work was intended to provide further independent and objective testing of existing analytical approaches, I chose to use obsidian as a "control" material with which to assess internal validity, because it has been used more than any other type of stone in the development of methods for debitage analysis. If the three approaches I was examining could not accurately identify the reduction strategy that produced the debitage with this best-case material, then I could safely assume that they were not valid analytical techniques, even at this basic, "internal" level. The results of this section on internal validity, therefore, refer only to the experiments I conducted using obsidian.

Of the three approaches, Magne's was the only one to correctly identify the reduction stages from which the obsidian flaking debris was derived in my experiments. While not *all* of the debitage in these assemblages was correctly assigned to its reduction stage, it did accurately classify most of "Early", or core reduction, debitage (including the bipolar, and large and medium prepared core debitage), much of the "Middle" stage samples, and it clearly pinpointed the "Late" stage soft hammer debitage, based on the highest percentage of one of the three flake classes. Since Magne did not supply a specific level of "fit" in terms of acceptable or expected values for Early, Middle, and Late flake class distributions, it was difficult to determine just how valid this approach actually was. Nevertheless, this approach is considered to be internally valid from the initial levels of analysis within the original types of experiments themselves.

In my experiments, the modified Sullivan and Rozen typology, however, was not able to distinguish between prepared core reduction and bifacial tool production assemblages. The obsidian debitage from prepared core reduction operations was dominated by small medial-distal fragments and complete flakes and did not follow the predictions proposed by Prentiss and Romanski (1989:92) for moderate amounts of these flake types and high percentages of nonorientable debris. The MSRT also failed to produce independent patterning for biface production debitage; in fact, the average Stage III assemblage showed close similarities to the core reduction flake profile, contrary to expectations. Bipolar flake profiles, however, did stand out noticeably in comparison to those created by other reduction techniques, since most of the assemblage was made up of medial-distal fragments (50% in the small size category and 17% in the medium). In fact, of the three approaches studied, this typology provided the clearest identification of obsidian bipolar debitage.

A mixture of success and failure marked the Mass Analysis results for obsidian. The flake count variables, for example, were not markedly different between the various manufacturing stages and reduction technologies, and the ratios thought to capture flake size variation also showed similar generalized flake trends for obsidian, with the only separation being the slight elevation of the Stage III assemblages. In addition, the relative frequencies of cortical flakes did not differ a great deal from one reduction strategy to the next, or at least they did not show any diagnostic patterns unless the average values (which eliminate much of the variability) were considered. On the other hand, the percentages of weight by size grade did at least show relative differences between Stage III biface, bipolar, and the other hard percussor debitage. Mean flake weights, as well, were useful since they indicated some patterning which is unique to each type or stage of reduction. Overall, while many of the Mass Analysis variables did not appear to be valid, there *were* some promising results that could be used to determine the technological origins of a flaking debris sample.

External Validity

In order to determine if the three methods could provide accurate inferences outside of the original experimental settings, I examined the debitage from the quartzite operations from the different perspectives of these approaches. I wanted to see if the flake patterns were the same as those for obsidian, or whether quartzite had its own debitage "signature" that would prevent any direct correlations or extrapolation from obsidian's results, patterns, and values. What I found was that there are a number of problems with the validity of these methods in an external sense.

Unlike the success I originally obtained from Magne's approach, the conclusions regarding the *quartzite* assemblages turned out to be inaccurate. Instead of providing clear distinctions between reduction stages and techniques, this method classified almost all of the quartzite debitage in the same manner (as "Early" core reduction debris). Since all the quartzite assemblages looked the same, this demonstrates the inability of this approach to be accurately extended to all flaked raw materials. One avenue that did show some promise was the platform flake/shatter ratio. While it is not possible to compare the results between different raw materials, this ratio did show distinctions between the quartzite reduction techniques in general; that is, this value was different for the hard hammer, soft hammer, and bipolar quartzite debitage assemblages.

On the other hand, when applying the MSRT to quartzite operations, problems were encountered because the debitage took on a generalized pattern for the two sizes of prepared cores as well as for biface reduction techniques. This standard pattern was primarily made up of small medial-distal fragments and nonorientable debris, with relatively small changes in the frequency of the other flake classes. Therefore, the independent patterning thought to differentiate between prepared core reduction and biface production did not exist for all types of lithic material. This means that the MSRT may not provide accurate technological identifications when applied to quartzite and other non-vitreous materials, although, again, some differences were evident, particularly in the bipolar profile which was dominated by a unique pattern of very high counts of small medial-distal and nonorientable fragments.

Most variables used in the interpretation of the Mass Analysis technique were also shown to be invalid for quartzite assemblages because they could not distinguish between the stages of biface manufacture or the types of core reduction. The results in this case were similar to those from the obsidian operations, with the most successful being the average cortex frequencies and the two weight variables (percentage weight by size grade, and mean flake weight), and the least successful being the flake count percentages. In addition, the ratio of G4: G1 to G3 flake counts could not differentiate between most quartzite operations either, although the bipolar reductions did produce flake count ratios that were quite unique.

Therefore, in a sense, the Mass Analysis approach was successful; however, the failure of some of the variables employed demonstrates the fact that there are many unexplored sources of variation in debitage assemblages. In general, I found this to be true of all the approaches studied, although in varying degrees for each one.

WHY WERE THE EXPECTATIONS NOT MET?

Since the differences between the obsidian and the quartzite assemblages are so extreme, and since most of these methods could not meet the expectations of the original researchers, it must be assumed that there are serious problems that either have to do with the manner in which the data was collected, or with the theoretical background (that is, the procedures simply do not measure what they are intended to measure because the theoretical constructs are somehow wrong) (Nance 1987:287).

Problems in collecting the data and applying the procedures

When considering the validity of any measuring instrument, it is vital to consider the effects of random and non-random (or systematic) error, as these can create artificial patterning in the data and can render the instrument or procedure both unreliable and invalid.

In my research, I found that one of the most important sources of error in the data collection phase came from having to rely on the ambiguous definitions of flake types or vague criteria for classifying certain kinds of flakes. This is especially noticeable in Magne's method which calls for the identification of biface thinning flakes and bipolar flakes, based on a few general characteristics for each class. However, as Rozen and Sullivan (1989a:170) have noted, by forcing the analyst to make a "best fit" judgment in order to place the flake in one category, this will obscure much of the potential information inherent in the collection. Despite this criticism, similar judgments have to be made when using their method as well. I learned about this problem when I had trouble classifying some quartzite flakes, since the coarse texture of this material does not always show the features used in the Sullivan and Rozen (1985) scheme. Again, with Magne's method, I noted a similar difficulty in quantifying flake scars, which are not always as obvious or distinct on some flakes (especially quartzite flakes). While errors like this can also occur throughout the application of the procedure, Magne (1985: 254) has claimed that this will only happen in low frequencies. The effects of split flakes and secondary multiple flaking on the platform area can also create a systematic error in Magne's classification, since it becomes necessary to classify these kinds of flakes as shatter rather than platform remnant-bearing flakes, since the striking platforms are not complete. Also, with secondary multiple flaking, additional dorsal flake scars are created which can change the classification of debitage to a "later" reduction stage, thus interfering with the correct systemic identification of an assemblage.

Another source of bias in Magne's and Ahler's methods is the removal of "usable" flakes from the debitage assemblages prior to analysis, in an effort to simulate flake-blank selection. However, the removal of this type of flake (for which the definitions are extremely vague) sets up a systematic error because the selection process is subjective. Furthermore, the rationale for choosing primarily large flakes may be unwarranted, since size was not always the deciding factor for a flake tool, at least in ethnographic accounts where the main criterion was edge angle (Prentiss 1993:77).

Yet another source of bias is found in Mass Analysis in the requirement for using comparative data as the basis for making inferences. In this respect, Ahler's approach is likely to generate invalid inferences, because all analysts who wish to apply their Mass Analysis results must rely on comparisons between their own data and the data published by Ahler; there is no other way to interpret one's results. As Beck and Jones (1989:246) have warned, "patterns observed are assumed to reflect archaeological patterns; it is rarely considered that there may be patterns that were created because of differences in analytic perception". More importantly, this type of error can then be *compounded* when further analysis and manipulation of the data takes place, particularly through statistical analyses. While this dilemma may never be completely eradicated in typological analyses, it can, in fact, be more closely controlled than it has in the past.

Theoretical Flaws and Weaknesses

Despite the numerous confounding variables and other problems associated with carrying out the three procedures, another possible explanation does exist for the lack of validity. Rather than being related directly to the practical aspects of analysis, this deals with inadequacies or inaccuracies in the theoretical building blocks with which these approaches are constructed. The amount of cortex cover, for example, is one variable that is thought to reflect stages of manufacture, even though it is probably misleading since cortical flakes can be encountered in all phases of tool production. Ahler's unquestioning use of this attribute may therefore generate invalid inferences regarding the systemic context of a debitage sample.

Although this specific theoretical flaw is particular to Mass Analysis, there are others that affect all three approaches with equal severity. As mentioned earlier, the stage concept of lithic reduction may well be an unsubstantiated and inflexible construct with little regard for situational needs or changes. Consequently, any approach that tries to place debitage components within strictly-defined stages or other technological boundaries may be unfounded, when it is considered how much variation there actually is in a real knapping operation.

Perhaps there are so many validity problems with these three approaches because they are so restricted in their methodologies. As a result, sometimes only the original experimental conditions can be interpreted with any degree of accuracy (see Ingbar et al 1989:123). This may also explain why some of Ahler's variables could not identify the reduction processes responsible for my experimental debitage assemblages. Since there was a great deal of variability both between the two raw material types and even within just one, his values and comparative data (which are based on the averaging of hundreds of experiments) may not be suitable for interpreting single cases. This is because the process of averaging tends to "dilute" the data into generalizations which can obscure the real range of variability that can be present in each individual case. I found this to be true in some of my own experiments, as well. The issue of variability is another area of lithic analysis that needs to be more fully considered when we interpret our data.

Construct Validity

The last possible reason for the discrepancies between the expectations and the experimental findings is that these approaches simply do not measure or identify the actions and strategies that they attempt to. As mentioned above, incorrect correlations between attributes and knapping procedures can create artificial patterning in the data which, in turn, can produce misleading theories on assemblage formation. However, sometimes the problem runs deeper than that, and we must re-evaluate the most basic constructs and assumptions that guide our reasoning.

In debitage analysis, one of these fundamental assumptions has always been that physical differences between raw material types will not have a significant effect on the patterning of flaking debris. This assumption is rather naïve, and it may well be responsible for producing unfounded conclusions on the meaning of lithic technological variation. While raw material factors are not the *only* ones influencing this, I have demonstrated that they do play a significant role in determining the patterns and distributions of flake classes.

I have shown that the knapping of coarser materials like quartzite requires its own unique reduction strategies and produces vastly different by-products than vitreous types of stone. For example, in comparison to obsidian, quartzite operations require much more force and generally rely on heavier percussors; they produce fewer and heavier flakes, and the distributions of debitage classes is unlike that of any other material yet tested. Furthermore, when this debitage is analysed according to methods tested only on obsidian or chert, the results are often a confusing array of indistinct or overlapping patterns and so-called diagnostic features. Consequently, the inferences drawn from such analysis cannot accurately characterize or identify the technological origins of the quartzite assemblage.

We need to thoroughly test not only our typologies but also our assumptions and theoretical constructs *before* we apply them to external situations which may be radically different in many ways. We also need to expand the range of knapping operations to include reduction techniques other than bifacial and core reductions, and involve a much broader array of raw material types that are truly representative of prehistoric technology. If we insist on relying on the results gained from only one or two lithic types, we run the risk of producing meaningless and invalid interpretations of the past.

By focusing solely on obsidian (or on chert, for that matter), lithic analysts may be simplifying some of their experiments, but they are only widening the rift between actualistic research and the archaeological record.

Recommendations

In the future, it would be beneficial to include petrological information on lithic raw materials when reporting both replicative *and* archaeological research, since this would ensure a higher level of consistency from one study to the next, and it would also enhance our understanding of how physical attributes can affect lithic debitage patterning. In addition, it is important, when conducting lithic analysis, to first sort debitage according to raw material type and to then analyse these samples separately. Polythetic attribute lists or flake type definitions should also be used whenever possible so as to avoid any typological confusion.

Furthermore, despite the strong indications of validity problems, my research does point out a number of promising aspects in the methods I evaluated, such as the success of Magne's platform flake/shatter ratio or the Mean Flake Weight variable employed in Mass Analysis; lithic analysts should take advantage of these potentially insightful measures. My findings still suggest that there is sometimes a certain degree of success in identifying gross differences in the flaking debris from various reduction techniques such as hard percussor and soft percussor flaking debris. Bipolar debitage, in particular, stood apart from the other techniques, especially when using the modified Sullivan and Rozen typology. Perhaps experimental research should continue to focus more on identifying these types of differences in flaking debris assemblages, rather than on trying to separate reduction *stages* (which, some researchers argue, do not even exist) into discrete, sequential debitage profiles. Patterning in debitage does exist, but we need to adjust our analytical procedures and interpretive devices to account for raw material differences and other sources of variability.

In this regard, I would recommend that both the theoretical and the practical aspects of experimental lithic analysis be tested for reliability and validity whenever possible, as this is the one way we can ensure that our assumptions and interpretations will be of value. We should also recognize the fact that more research needs to be carried out, using larger and more varied samples, before we can assume that a method of analysis will work. While my experiments may have contributed to our understanding of the variability that can be present in lithic debitage, they are only a starting point from which to proceed.

REFERENCES CITED

Ahler, Stanley A.

1989a Experimental Knapping with KRF and Midcontinent Cherts: Overview and Applications. In *Experiments in Lithic Technology*, edited by Daniel Amick and Raymond Mauldin, pp.199-234. BAR International Series No. 528, British Archaeological Reports, Oxford.

1989b Mass Analysis of Flaking Debris: Studying the Forest rather than the tree. In *Alternative Approaches to Lithic Analysis*, edited by Donald Henry and George Odell, pp. 85-118. Archaeological Papers of the American Anthropological Association Number 1.

Ahler, Stanley A. and Julieann VanNest

1984 Temporal Change in Knife River Flint Reduction Strategies, In *Lithic Resource Procurement: Proceedings from the Second Conference on Prehistoric Chert Exploitation*, edited by Susan C.. Vehik, pp.183-198. Center for Archaeological Investigations, Occasional Paper No. 4. Southern Illinois University, Carbondale.

Amick, Daniel S. and Raymond P. Mauldin

1989 Comments on Sullivan and Rozen's "Debitage analysis and archaeological interpretation". *American Antiquity* 54(1): 166-168.

Amick, Daniel S., Raymond P. Mauldin and Lewis R. Binford

1989 The Potential of Experiments in Lithic Technology. In *Experiments in Lithic Technology*, edited by Daniel Amick and Raymond Mauldin, pp. 1-14. BAR International Series No.528, British Archaeological Reports, Oxford.

Baumler, Mark F,

1985 On the interpretation of chipping debris concentrations in the archaeological record. *Lithic Technology* 14(3):120-125.

Baumler, Mark F. and Christian E. Downum

1989 Between Micro and Macro: A Study in the Interpretation of Small-Sized Lithic Debitage. In *Experiments in Lithic Technology*, edited by Daniel Amick and Raymond Mauldin, pp.101-116. BAR International Series No. 528, British Archaeological Reports, Oxford.

Beck, Charlotte and George T. Jones

1989 Bias and Archaeological Classification. *American Antiquity* 54(2): 244-262.

1990 Toolstone selection and lithic technology in early Great Basin prehistory. *Journal of Field Archaeology* 17:283-299.

Behm, Jeffery A.
1983 Flake concentrations: Distinguishing between flintworking activity areas and secondary deposits. *Lithic Technology* 12(1):9-16.

Behm, Jeffery A. and Alaric Faulkner
1974 Hixton Quartzite: Experiments in Heat Treatment. *The Wisconsin Archeologist* 55(4): 271-276.

Best, Myron G.
1982 *Igneous and Metamorphic Petrology.* W.H. Freeman & Co., San Francisco.

Bies, Michael T. and Eugene J. Romanski
1988 An intensive cultural resource inventory of the Hamilton Desert Land Entry (W-84574), Big Horn County, Wyoming. *Worland District BLM Cultural Resource Report.* Worland, Wyoming.

Binford, Lewis R. and James F. O'Connell
1984 An Alyawara day: the stone quarry . *Journal of Anthropological Research* 40(3):406-432.

Binford, Lewis R. and G.I. Quimby
1963 Indian sites and chipped stone materials in the northern Lake Michigan region. *Fieldiana Anthropology* 36:277-307.

Boksenbaum, Martin W.
1980 Basic Mesoamerican Stone-Working: Nodule Smashing? *Lithic Technology* 9(1): 12-25.

Bordes, Francois
1961 *Typologie du Paléolithique Ancien et Moyen.* Publications de préhistoire de l'Université de Bordeaux, Memoire 1.

Bouey, Paul D.
1983 Volume as a Measure of Lithic Debitage. *Lithic Technology* 12(2):39-42.

Bowen, Robert
1984 *Geology in Engineering.* Elsevier Applied Science Publishers, London.

Bradley, Bruce
1974 Comments on the Lithic Technology of the Casper Site Materials. pp. 191-197 In *The Casper Site: A Hell Gap Bison Kill on the High Plains*, edited by George Frison, pp.191-197. Academic Press, New York.

Bridgland, David R.
1988 The Quaternary Derivation of Quartzites used by Palaeolithic Man in the Thames Basin for Tool Manufacture. In *Non-Flint Stone Tools and the Palaeolithic Occupation of Britain,* edited by R.J. MacRae and N. Moloney, pp.187-198. BAR British Series No. 189, British Archaeological Reports, Oxford.

Brisland, Richard T.W.
1992 *Camas Processing or Upland Hunting: An Interpretation of Lithic Scatters at High Prairie.* Unpublished M.A. Thesis, Department of Archaeology, University of Calgary.

Brothwell, Don
1983 Petrology and Archaeology: an introduction. In *The Petrology of Archaeological Artefacts* , edited by D.R.C. Kempe and Anthony P. Harvey, pp.1-25. Clarendon Press, Oxford.

Brown, Kenneth L. , Marie E. Brown, and Ned H. Hanenberger
1982 *Prehistoric Stone Tools of South Dakota: A Guide.* Special Publication of the South Dakota Archaeological Society, No. 6. University of South Dakota Archaeology Laboratory, Vermilion.

Burton, John
1980 Making sense of waste flakes: new methods for investigating the technology and economics behind chipped stone assemblages. *Journal of Archaeological Science* 7:131-148.

Cahen, D. and F. Van Noten
1971 Stone Age Typology: Another Approach. *Current Anthropology* 12(2):211-214.

Callahan, Errett
1979 The Basics of Biface Knapping in the Eastern Fluted Point Tradition: A Manual for Flintknappers and Lithic Analysts. *Archaeology of Eastern North America* 7(1): 1-213.

Cann, J.R.
1983 Petrology of obsidian artefacts. In *The Petrology of Archaeological Artefacts* , edited by D.R.C. Kempe and Anthony P. Harvey, pp. 227-254. Clarendon Press, Oxford.

Carmichael, Robert S.
1989 *Practical Handbook of Physical Properties of Rocks and Minerals.* CRC Press, Boca Raton.

Carmines, Edward G. and Richard A. Zeller
1979 *Reliability and Validity Assessment.* SAGE Publications, Beverly Hills.

Chandler, Susan M. and John A. Ware
1976 The identification of technological variability through experimental replication and empirical multivariate analysis. *Newsletter of Lithic Technology* 5(3): 24-26.

Clark, John E.
1986 Another look at small debitage and microdebitage. *Lithic Technology* 15(1):21-33.

Clough, T.H. McK. and A.R. Woolley
1985 Petrography and stone implements. *World Archaeology* 17(1):92-100.

Coe, Joffre
1975 Comment. In *Lithic Technology: Making and Using Stone Tools*, edited by Earl Swanson, pp. 229-232. Mouton Publishing, The Hague.

Cotterell, Brian and Johan Kamminga
1986 Finials on Stone Tools. *Journal of Archaeological Science* 13: 451-461.

1987 The Formation of Flakes. *American Antiquity* 52(4): 675-708.

Coulson, Sheila and Birgitte Skar
1990 Introduction to The Second Flint Alternatives Workshop. *Særtrykk* (Årbok, Universitetets Oldsaksamling) 1989/1991: 71.

Crabtree, Don E.
1967 Notes on Experiments in Flintknapping: 3. The Flintknapper's Raw Materials. *Tebiwa* 10: 8- 25.

1972 *An introduction to flintworking.* Occasional Papers of the Idaho State University Museum No. 28, Pocatello.

1975 Comments on Lithic Technology and Experimental Archaeology. In *Lithic Technology: Making and Using Stone Tools*, edited by Earl Swanson, pp.105-114. Mouton Publishers, The Hague.

Cummins, W.A.
1983 Petrology of stone axes and tools. In *The Petrology of Archaeological Artefacts*, edited by D.R.C. Kempe and Anthony P. Harvey,pp.171-226. Clarendon Press, Oxford.

Davis, Dave D.
1978 Lithic assemblage variability in relation to early hominid subsistence strategies at Olduvai Gorge. In *Lithics and Subsistence: The analysis of stone tool use in prehistoric economies* , edited by Dave D. Davis, pp. 355-86. Publications in Anthropology No. 20, Vanderbilt University, Nashville.

Dibble, Harold L.
1985 Technological Aspects of Flake Variation: A Comparison of Experimental and Prehistoric Flake Production. *American Archaeology* 5(3):236-240.

1987 Reduction Sequences in the Manufacture of Mousterian Implements of France. In *The Pleistocene Old World*, edited by Olga Soffer, pp. 33-45. Plenum Press, New York.

Dibble, Harold L. and John C. Whittaker
1981 New experimental evidence on the relation between percussion flaking and flake variation. *Journal of Archaeological Science* 8:283-296.

Duncan, Neil
1969 *Engineering Geology and Rock Mechanics, Volume I.* CRC Press (International Scientific Series), London.

Ebright, Carol A.
1978 Comment on "Quartzite qua lithic material in archaeology" by H. Wolcott Toll. *Plains Anthropologist* 24(84, part 1): 169-170.

1987 Quartzite Petrography and its Implications for Prehistoric Use and Archaeological Analysis. *Archaeology of Eastern North America* 15: 29-45.

Ellis, Christopher J. and Jonathan C. Lothrop
1989 *Eastern Paleoindian Lithic Resource Use.* Westview Press, Boulder.

Ensor, H. Blaine and Erwin Roemer, Jr.
1989 Comments on Sullivan and Rozen's Debitage Analysis and Archaeological Interpretation. *American Antiquity* 54(1): 175-178.

Fish, Paul R.
1978 Consistency in archaeological measurement and classification: a pilot study. *American Antiquity* 43(1): 86-89.

Fladmark, Knut R.
1982 Microdebitage Analysis: Initial Considerations. *Journal of Archaeological Science* 9:205-220.

Flenniken, J. Jeffrey

1981 *Replicative Systems Analysis: a model applied to the vein quartz artifacts from the Hoko River site.* Laboratory of Anthropology Reports of Investigations No.59, Washington State University, Pullman.

Flenniken, J. Jeffrey and E.G. Garrison

1975 Thermally altered novaculite and stone tool manufacturing techniques. *Journal of Field Archaeology* 2:125-131.

Francis, Julie

1983 *Lithic Procurement Strategies in the Big Horn Basin and Mountains of North-Central Wyoming*. Unpublished Ph.D. Dissertation, Department of Anthropology, Arizona State University.

Frison, George

1974 Archaeology of the Casper Site. In *The Casper Site: A Hell Gap Bison Kill on the High Plains*, edited by George Frison, pp.1-111. Academic Press, New York.

1982 Raw Stone Flaking Material Sources. In *The Agate Basin Site: A record of the Paleo-Indian Occupation of the Northwestern High Plains*, edited by George E. Frison and Dennis J. Stanford, pp. 173-178. Academic Press, New York.

Frison, George and Bruce Bradley

1980 *Folsom Tools and Technology at the Hanson Site, Wyoming*. University of New Mexico, Albuquerque.

Goodman, Mary Ellen

1944 The Physical Properties of Stone Tool Materials. *American Antiquity* 4: 415-433.

Gould, Richard

1978 Beyond Analogy in Ethnoarchaeology. In *Explorations in Ethnoarchaeology*, edited by Richard Gould, pp. 249-293. University of New Mexico Press, Albuquerque.

Greiser, Sally Thompson

1985 Predictive Models of Hunter-Gatherer Subsistence and Settlement Strategies on the Central High Plains. *Plains Anthropologist* 30 (110, part 2)/ Memoir 20 Plains Anthropologist.

Greiser, Sally T. and Payson D. Sheets

1979 Raw Material as a Functional Variable in Use-Wear Studies. pp. 289-296 In *Lithic Use-Wear Analysis*, edited by Brian Hayden, pp. 289-296. Academic Press, New York.

Gunn, Joel

1975 Idiosyncratic Behavior in Chipping Style: Some Hypotheses and Preliminary Analysis. In *Lithic Technology: Making and Using Stone Tools*, edited by Earl Swanson, pp.35-61. Mouton Publishers, The Hague.

Hay, Conron A. and Robert Rogers

1978 A method for determining polyhedral core sizes from prismatic blade debitage. *Lithic Technology* 7(3): 46-52.

Hayden, Brian

1980 Confusion in the Bipolar World: Bashed Pebbles and Splintered Pieces. *Lithic Technology* 9(1):2-7.

Hayden, Brian, Nora Franco and Jim Spafford

1993 "Evaluating Lithic Strategies and Design Criteria". Paper presented at The Conference on Theory in Lithics, Tulsa, Oklahoma.

Hayden, Brian and W. Karl Hutchings

1989 Whither the Billet Flake? In *Experiments in Lithic Technology*, edited by Daniel S. Amick and Raymond P. Mauldin, pp. 235-257. BAR International Series No. 528, British Archaeological Reports, Oxford.

Hayden, Brian and Johan Kamminga

1979 An Introduction to Use-Wear. In *Lithic Use-Wear Analysis*, edited by Brian Hayden, pp. 1-14. Academic Press, New York.

Henry, Don O., C. Vance Haynes and Bruce Bradley

1976 Quantitative Variations in Flaked Stone Debitage. *Plains Anthropologist* 21:57-61.

Hiscock, Peter

1986 Technological change in the Hunter River valley and the interpretation of late Holocene change in Australia. *Archaeology in Oceania* 21(1): 40-50.

Holm, Lena

1990 The Use of Stone in the Rana-Tärna region, North Scandinavia. *Særtrykk* (Årbok, Universitetets Oldsaksamling)1989/1990:85-91.

Hull, Kathleen

1987 Identification of Cultural Site Formation Processes through Micro-Debitage Analysis. *American Antiquity* 52: 772-783.

Ingbar, Eric E., Mary Lou Larson and Bruce A. Bradley

1989 A Nontypological Approach to Debitage Analysis. In *Experiments in Lithic Technology*, edited by Daniel Amick and Raymond Mauldin, pp. 117-136. BAR International Series 528, British Archaeological Reports, Oxford.

Inskeep, R.R.

1988 Africa: Everything but Flint. In *Non-Flint Stone Tools and the Palaeolithic Occupation of Britain*, edited by R.J. MacRae and N. Moloney, pp. 235-242. BAR British Series No. 189,British Archaeological Reports, Oxford.

Jelinek, Arthur J., Bruce Bradley and Bruce Huckell

1971 The Production of Secondary Multiple Flakes. *American Antiquity* 36(2): 198-200.

Johnson, L. Lewis

1975 Graph Theoretical Analysis of Lithic Tools from Northern Chile. In *Lithic Technology: Making and Using Stone Tools*, edited by Earl Swanson, pp. 63-96. Mouton Publishing, The Hague.

Jones, Kevin L.

1984 Lithic waste flakes as a measure of cultural affinity: a New Zealand case study. *Lithic Technology*13(3): 71-83.

Jones, Peter Rasmussen

1978 Effects of Raw Materials on Biface Manufacture. *Science* 204: 835-836.

Kamminga, Johan

1982 *Over the Edge: Functional Analysis of Australian Stone Tools*. Occasional Paper in Anthropology No. 12. University of Queensland Printery, St. Lucia.

Kelly, Robert L.

1988 The Three Sides of a Biface. *American Antiquity* 53(4): 717-734.

Kempe, D.R.C. and J.A. Templeman

1983 Techniques. In *The Petrology of Archaeological Artefacts, edited* by D.R.C. Kempe and Anthony P. Harvey, pp. 26-52. Clarendon Press, Oxford.

Kleindienst, Maxine

1979 Discussion: Raw Materials. In *Lithic Use-Wear Analysis*, edited by Brian Hayden, pp. 297-299.

Knudson, Ruthann
1978 Experimental Lithicology: Method and Theory. *Lithic Technology* 7(3):44-46.

1983 *Organizational Variability in Late Paleo-Indian Assemblages*. Reports of Investigations, No. 60. Laboratory of Anthropology, Washington State University, Pullman.

Kuhn, Steven L.
1989 Hunter-Gatherer Foraging Organization and Strategies of Artifact Replacement and Discard. In *Experiments in Lithic Technology,* edited by Daniel Amick and Raymond Mauldin, pp. 33-48. BAR International Series No. 528, British Archaeological Reports, Oxford.

1991 "Unpacking" Reduction: Lithic Raw Material Economy in the Mousterian of West-Central Italy. *Journal of Anthropological Archaeology* 10: 76-106.

Lee, Thomas E.
1954 The First Sheguiandah Expedition, Manitoulin Island, Ontario. *American Antiquity* 20(2):101-111.

1955 The Second Sheguiandah Expedition, Manitoulin Island, Ontario. *American Antiquity* 21(1): 63-71.

Luedtke, Barbara E.
1976 *Lithic Material Distributions and Interaction Patterns During the Late Woodland Period in Michigan, Volumes I and II.* Ph.D. Dissertation. University Microfilms International, Ann Arbor.

1992 *An Archaeologist's Guide to Chert and Flint.* Archaeological Research Tools 7. Institute of Archaeology, University of California, Los Angeles.

MacRae, R. J.
1988a The Palaeolithic of the Upper Thames and its Quartzite Implements. In *Non-Flint Stone Tools and the Paleaolithic Occupation of Britain*, edited by R.J. MacRae and N. Moloney, pp. 123-154. BAR British Series No. 189, British Archaeological Reports, Oxford.

1988b The Palaeolithic Artefacts of Warwickshire. In *Non-Flint Stone Tools and the Palaeolithic Occupation of Britain*, edited by R.J. MacRae and N. Moloney, pp. 95-102. BAR British Series No. 189, British Archaeological Reports, Oxford.

MacRae, R.J. and N. Moloney
1988 *Non-Flint Stone Tools and the Paleolithic Occupation of Britain.* BAR British Series No. 188, British Archaeological Reports, Oxford.

Magne, Martin P.R.
1985 *Lithics and Livelihood: Stone tool technologies of central and southern British Columbia.* National Museum of Man, Mercury Series, Archaeological Survey of Canada Paper No. 133.

1989 Lithic Reduction Stages and Assemblage Formation. In *Experiments in Lithic Technology*, edited by Daniel Amick and Raymond Mauldin, pp. 15-32. BAR International Series No. 528, British Archaeological Reports, Oxford.

Mandeville, M.D.
1973 A consideration of the thermal pretreatment of chert. *Plains Anthropologist* 18: 177-202.

Mauldin, Raymond P. and Daniel S. Amick
1989 Investigating Patterning in Debitage from Experimental Bifacial Core Reduction. In *Experiments in Lithic Technology,* edited by Daniel Amick and Raymond Mauldin, pp. 67-88. BAR International Series No. 528, British Archaeological Reports, Oxford.

Minor, Rick and Kathryn Anne Toepel
1982 The Halverson Site: A Late Prehistoric campsite in the Upper Willamette Valley, Oregon. *Tebiwa* 19:1-14.

Moloney, N.
1988 Experimental Biface Manufacture using Non-Flint Lithic Materials. In *Non-Flint Stone Tools and the Palaeolithic Occupation of Britain*, edited by R.J. MacRae and N. Moloney, pp. 49-65. BAR British Series No. 189, British Archaeological Reports, Oxford.

Moloney, N., C. Bergman, M. Newcomer, and F. Wenban-Smith
1988 Experimental Replication of Bifacial Implements using Bunter Quartzite Pebbles. In *Non-Flint Stone Tools and the Palaeolithic Occupation of Britain*, edited by R.J. MacRae and N. Moloney, pp. 25-47. BAR British Series No. 189, British Archaeological Reports, Oxford.

Moore, David S.
1991 *Statistics: Concepts and Controversies.* W.H. Freeman and Company, New York.

Nance, J.D.
1987 Reliability, Validity, and Quantitative Methods in Archaeology. In *Quantitative Research in Archaeology*, edited by M. Aldenderfer, pp. 244-293. Sage, Beverly Hills.

Nance, J.D. and B.F. Ball
1986 No Surprises? The Reliability and Validity of Test Pit Sampling. *American Antiquity* 54(2): 457-483.

Newcomer, M.H.
1975 Punch Technique"and Upper Paleolithic Blades. In *Lithic Technology: Making and Using Stone Tools* , edited by Earl Swanson, pp. 97-104. Mouton Publishing, The Hague.

1973 Some Quantitative Experiments in Handaxe Manufacture. *World Archaeology* 3: 85-93.

Newcomer, M.H. and C. Karlin
1987 Flint Chips from Pincevent. In *The Human Uses of Flint and Chert,* edited by G.de G. Sieveking and M.H. Newcomer, pp. 33-36. Cambridge University Press, Cambridge.

Nichols, Jacqueline and Diane J. Allstadt
1978 Hinge fracture rates of novice flintknappers. *Lithic Technology* 7(1): 1-2.

Odell, George H.
1980 Butchering with stone tools: some experimental results. *Lithic Technology* 9(2): 39-48.

1989 Experiments in Lithic Reduction. In *Experiments in Lithic Technology,* edited by Daniel Amick and Raymond Mauldin, pp. 163-198. BAR International Series No. 528, British Archaeological Reports, Oxford.

Ohel, Milla
1978 "Clactonian Flaking" and Primary Flaking: Some initial considerations. *Lithic Technology* 7(1): 23-28.

Patterson, L.W.
1982 The Importance of Flake Size Distribution. *Contract Abstracts and CRM Archeology* 3(1):70-72.

1982 Replication and classification of large size lithic debitage. *Lithic Technology* 11(3): 50-58.

1990 Characteristics of Bifacial-Reduction Flake-Size Distribution. *American Antiquity* 55(3): 550-558.

Patterson, L.W. and J.B. Sollberger
1978 Replication and classification of small size lithic debitage. *Plains Anthropologist* 23: 103-112.

Pearl, Richard M.
1956 *Rocks and Minerals*. Barnes and Noble, New York.

Pettijohn, F. J.
1957 *Sedimentary Rocks*. Harper and Row, New York.

Pokotylo, David L. and Christopher C. Hanks
1989 Measuring Assemblage Variability in Curated Lithic Technologies: A Case Study from the Mackenzie Mountains, Northwest Territories. In *Experiments in Lithic Technology*, edited by Daniel Amick and Raymond Mauldin, pp. 49-66. BAR International Series No. 528, British Archaeological Reports, Oxford.

Pond, Alonzo W.
1930 *Primitive Methods of Working Stone, Based on Experiments of Halvor Skavlem.* Logan Museum Bulletin 3, Beloit College.

Prentiss, William C.
1991 *Proposal for Archaeological Investigations at 48BH932*. Paper submitted to the Bureau of Land Management, Worland, Wyoming.

1993 *Hunter-Gatherer Economics and the Formation of a Housepit Floor Lithic Assemblage*. Unpublished Ph.D. Dissertation, Simon Fraser University, Burnaby, British Columbia.

Prentiss, William C. and Eugene J. Romanski
1989 Experimental Evaluation of Sullivan and Rozen's Debitage Typology. In *Experiments in Lithic Technology*, edited by Daniel Amick and Raymond Mauldin, pp. 89-100. BAR International Series No. 528, British Archaeological Reports, Oxford.

Prentiss, William C., E.J. Romanski and M.L. Douhit
1988 Hunter-Gatherer Land Use and Lithic Procurement in the Central Big Horn Basin, Wyoming. *The Wyoming Archaeologist* 31(1-2): 33-53.

Preston, Neil J.
1988 The Triassic Bunter Pebble Beds as a Source for Non-Flint Artefacts in the Upper Thames Basin. In *Non-Flint Stone Tools and the Palaeolithic Occupation of Britain*, edited by R.J. MacRae and N. Moloney, pp. 199-216. BAR British Series No. 189, British Archaeological Reports, Oxford.

Raab, L. Mark, Robert F. Cande, and David W. Stahle

1979 Debitage graphs and archaic settlement patterns in the Arkansas Ozarks. *Midcontinental Journal of Archaeology* 4(2):167-182.

Ranere, Anthony J.

1975 Toolmaking and Tool Use among the Preceramic Peoples of Panama. In *Lithic Technology: Making and Using Stone Tools*, edited by Earl Swanson, pp. 173-209. Mouton, The Hague.

Rankama, Tuija

1990 Quartzite at Utsjoki Ala-Jalve: The frame of a case study. *Særtrykk* (Årbok, Universitetets Oldsaksamling) 1989/1990:103-117.

Renfrew, Colin, J.R. Cann, and J.E. Dixon

1973 Obsidian and Trade in the Near East. In *In Search of Man: Readings in Archaeology*, edited by Ernestene L. Green, pp. 251-265. Little, Brown and Company, Boston.

Rozen, Kenneth C. and Alan P. Sullivan

1989a Measurement, Method, and Meaning in Lithic Analysis: Problems with Amick and Mauldin's Middle-Range Approach. *American Antiquity* 54(1): 169-175.

1989b The Nature of Lithic Reduction and Lithic Analysis: Stage Typologies Revisited. *American Antiquity* 54(1): 179-184.

Runnels, Curtis

1985 Lithic Studies: Some Theoretical Considerations. *Lithic Technology* 14(3): 100-106.

Schiffer, Michael B.

1976 *Behavioral Archaeology*. Academic Press, New York.

Semenov, S.A.

1964 *Prehistoric Technology*. Translated by M.W. Thompson. Redwood Press, Bath, Great Britain.

Shelley, Phillip H.

1990 Variation in lithic assemblages: An experiment. *Journal of Field Archaeology* 17(2):187-193.

Shott, Michael J.

1989 Bipolar Industries: Ethnographic evidence and archaeological implications. *North American Archaeologist* 10(1): 1-24.

Sieveking, G. de G.
1980 Experimental Flake Scatter-Patterns: a New Interpretive Technique. *Journal of Field Archaeology* 7:345-352.

Simpson, David N.
1990 The Impact of Post-Manufacture Alterations on Stone Tool Typology. *Særtrykk* (Årbok, Universitetets Oldsaksamling) 1989/1990:119-128.

Speth, John D.
1972 Mechanical Basis of Percussion Flaking. *American Antiquity* 37(1): 34-60.

1975 Miscellaneous Studies in Hard-Hammer Percussion Flaking: The Effects of Oblique Impact. *American Antiquity* 40(2): 203-207.

1981 The Role of Platform Angle and Core Size in Hard-Hammer Percussion Flaking. *Lithic Technology* 10(1): 16-21.

Spry, Alan
1969 *Metamorphic Textures*. Pergamon Press, Oxford.

Stafford, C. Russell and Barbara D. Stafford
1979 Some issues concerning the design of lithic experiments. *Lithic Technology* 8(2):21-24.

Stahle, David W. and James E. Dunn
1982 An analysis and application of the size distribution of waste flakes from the manufacture of bifacial stone tools. *World Archaeology* 14 (1): 84-97.

Stevenson, Christopher, Leslie Shaw and Claudio Cristino
1984 Obsidian Procurement and Consumption on Easter Island. *Archaeology of Oceania* 19(3): 120-124.

Strauss, Alan E.
1989 Narragansett Basin Argillite: Lithology, Chronology, and Prehistoric Tool Manufacture. *North American Archaeologist* 10(1): 25-37.

Strauss, Lawrence Guy
1980 The Role of Raw Materials in Lithic Assemblage Variability. *Lithic Technology* 9(3): 68-72.

Sullivan, Alan P. III
1987 Probing the Sources of Variability: A Regional Case Study near the Homolovi Ruins, Arizona. *North American Archaeologist* 8(1):41-71.

Sullivan, Alan P. III and Kenneth C. Rozen
1985 Debitage Analysis and Archaeological Interpretation. *American Antiquity* 50(4): 755-779.

Taçon, Paul S.C.
1991 The Power of Stone: symbolic aspects of stone tool use and tool development in western Arnhem Land, Australia. *Antiquity* 65: 192-207.

Taylor, R.E.
1976 Preface. In *Advances in Obsidian Glass Studies: Archaeological and Geochemical Perspectives*, edited by R.E. Taylor, pp. vii-viii. Noyes Press, Park Ridge.

Thomas, David Hurst
1986 Contemporary hunter-gatherer archaeology in America. In *American Archaeology Past and Future*, edited by D. Meltzer, L. Fowler, and J. Sabloff, pp. 237-276. Smithsonian Institution Press, Washington, D.C..

Toll, H. Wolcott
1978 Quartzite qua lithic material in archaeology: qualities and quandaries with special reference to use-wear. *Plains Anthropologist* 22(79): 47-67.

Tomka, Steven A.
1989 Differentiating Lithic Reduction Techniques: An Experimental Approach. In *Experiments in Lithic Technology*, edited by Daniel Amick and Raymond Mauldin, pp. 137-162. BAR International Series No. 528, British Archaeological Reports, Oxford.

Toth, Nicholas
1985 The Oldowan Reassessed: A close look at early stone artifacts. *Journal of Archaeological Science* 12:101-120.

Touloukian, Y.S., W.R. Judd and R.F. Roy
1981 *Physical Properties of Rocks and Minerals.* McGraw-Hill Cindas Data Series on Material Properties Volume II-2. McGraw-Hill, New York.

Tringham, Ruth
1978 Experimentation, Ethnoarchaeology, and the Leapfrogs in Archaeological Methodology. In *Explorations in Ethnoarchaeology*, edited by Richard Gould, pp. 169-200. University of New Mexico Press, Albuquerque.

Tuohy, Donald R.
1987 A Comparison of Pressure and Percussion Debitage from a Crabtree Obsidian Stoneworking Demonstration. *Tebiwa* 23: 23-30.

Vehik, Susan C.
1984 *Lithic Resource Procurement: Proceedings from the Second Conference on Prehistoric Chert Exploitation.* Center for Archaeological Investigations Occasional Paper No.4. Southern Illinois University, Carbondale.

Watson, Alison
1983 Chipped-Stone Artifacts from Servia in Greek Macedonia. *Journal of Field Archaeology* 10:120-123.

Wedel, Waldo
1978 *Prehistoric Man on the Great Plains.* University of Oklahoma Press, Norman.

Whitehead, P.F.
1988 Lower Palaeolithic Artefacts from the Lower Valley of the Warwickshire Avon. In *Non-Flint Stone Tools and the Palaeolithic Occupation of Britain*, edited by R.J. MacRae and N. Moloney, pp. 103-121. BAR British Series No. 189, British

Williams, Howel, Francis J. Turner and Charles M. Gilbert
1982 *Petrography: An introduction to the study of rocks in thin sections.* W.H. Freeman and Company, San Francisco.

Wymer, John
1988 A Reassessment of the Geographical Range of Lower Palaeolithic Activity in Britain. In *Non-Flint Stone Tools and the Palaeolithic Occupation of Britain*, edited by R.J. MacRae and N. Moloney, pp. 11-23. BAR British Series No. 189, British Archaeological Reports, Oxford.

Wynn, Thomas and Forrest Tierson
1990 Regional Comparison of the Shapes of Later Acheulian Handaxes. *American Anthropologist* 92(1): 73-84.

Yerkes, Richard W. and P. Nick Kardulias
1993 Recent Developments in the Analysis of Lithic Artifacts. *Journal of Archaeological Research* 1(2): 89-119.

APPENDIX:
Raw Data

FLAKE COUNTS
according to the Modified Sullivan and Rozen Typology

Legend:

CF = Complete Flake
PF = Proximal Flake
MD = Medial-Distal Flake
SF = Split Flake
NO = Nonorientable Flake

XL = Extra Large
L = Large
M = Medium
S = Small

QUARTZITE EXPERIMENTS

ASSEMBLAGE #2 -- BIPOLAR CORE REDUCTION

	XL	L	M	S	Row Totals
CF	2	3	0	0	5
PF	0	0	0	0	0
MD	0	7	12	38	57
SF	0	0	2	0	2
NO	0	0	6	29	35
Column Totals	2	10	20	67	99

ASSEMBLAGE # 12 -- BIPOLAR CORE REDUCTION

	XL	L	M	S	Row Totals
CF	0	2	1	3	6
PF	0	3	4	4	11
MD	0	3	5	37	45
SF	0	3	6	15	24
NO	0	0	4	26	30
Column Totals	0	11	20	85	116

ASSEMBLAGE # 20 -- MEDIUM-SIZED PREPARED CORE REDUCTION

	XL	L	M	S	Row Totals
CF	0	6	5	21	32
PF	0	1	6	12	19
MD	0	1	11	29	41
SF	0	5	8	6	19
NO	0	0	3	19	22
Column Totals	0	13	33	87	133

ASSEMBLAGE #3 -- MEDIUM-SIZED PREPARED CORE REDUCTION

	XL	L	M	S	Row Totals
CF	0	3	7	6	16
PF	0	7	7	11	25
MD	0	0	4	16	20
SF	0	1	5	6	12
NO	0	0	5	18	23
Column Totals	0	11	28	57	96

ASSEMBLAGE # 7 -- LARGE PREPARED CORE REDUCTION

	XL	L	M	S	Row Totals
CF	5	5	3	25	38
PF	0	6	4	19	29
MD	0	5	9	32	46
SF	0	4	11	8	23
NO	0	0	13	55	68
Column Totals	5	20	40	139	204

ASSEMBLAGE #5 -- LARGE PREPARED CORE REDUCTION

	XL	L	M	S	Row Totals
CF	3	5	4	16	28
PF	0	3	8	22	33
MD	0	7	13	40	60
SF	0	9	20	19	48
NO	0	0	15	61	76
Column Totals	3	24	60	158	239

ASSEMBLAGE # 8A -- STAGE II BIFACE REDUCTION (hard hammer)

	XL	L	M	S	Row Totals
CF	0	7	6	9	22
PF	0	2	7	16	25
MD	0	10	11	35	56
SF	0	3	11	19	33
NO	0	0	16	31	47
Column Totals	0	22	51	110	183

ASSEMBLAGE # 6A --STAGE II BIFACE REDUCTION (hard hammer)

	XL	L	M	S	Row Totals
CF	0	3	0	6	9
PF	0	2	8	8	18
MD	0	6	11	27	44
SF	0	2	4	0	6
NO	0	4	9	42	55
Column Totals	0	17	32	83	132

ASSEMBLAGE # 6B -- STAGE III BIFACE REDUCTION (soft hammer)

	XL	L	M	S	Row Totals
CF	0	0	4	10	14
PF	0	1	9	5	15
MD	0	1	11	14	26
SF	0	0	2	4	6
NO	0	0	1	8	9
Column Totals	0	2	27	41	70

ASSEMBLAGE #8B -- STAGE III BIFACE REDUCTION (soft hammer)

	XL	L	M	S	Row Totals
CF	0	12	17	20	49
PF	0	2	8	31	41
MD	0	2	30	55	87
SF	0	0	2	12	14
NO	0	0	6	52	58
Row Totals	0	16	63	170	249

OBSIDIAN EXPERIMENTS

ASSEMBLAGE # 13 -- BIPOLAR CORE REDUCTION

	XL	L	M	S	Row Totals
CF	1	5	5	7	18
PF	0	2	4	7	13
MD	0	8	21	74	103
SF	0	0	7	3	10
NO	0	0	3	13	16
Column Totals	1	15	40	104	160

ASSEMBLAGE # 14 -- BIPOLAR CORE REDUCTION

	XL	L	M	S	Row Totals
CF	0	2	1	2	5
PF	0	2	2	4	8
MD	0	2	20	51	73
SF	0	3	2	2	7
NO	0	0	1	2	3
Column Totals	0	9	26	61	96

ASSEMBLAGE # 11 -- MEDIUM-SIZED PREPARED CORE REDUCTION

	XL	L	M	S	Row Totals
CF	1	2	1	3	6
PF	0	3	4	4	11
MD	0	3	5	37	45
SF	0	3	6	15	24
NO	0	0	4	26	30
Column Totals	1	11	20	85	116

ASSEMBLAGE # 19 -- MEDIUM-SIZED PREPARED CORE REDUCTION

	XL	L	M	S	Row Totals
CF	0	20	6	19	45
PF	1	4	3	2	10
MD	0	7	13	45	65
SF	0	0	0	1	1
NO	0	0	1	2	3
Column Totals	1	31	23	69	123

ASSEMBLAGE # 17 -- LARGE PREPARED CORE REDUCTION

	XL	L	M	S	Row Totals
CF	1	6	5	38	50
PF	1	7	4	7	19
MD	0	9	9	46	64
SF	0	6	1	3	10
NO	0	0	3	6	9
Column Totals	2	28	22	100	152

ASSEMBLAGE # 18 -- LARGE PREPARED CORE REDUCTION

	XL	L	M	S	Row Totals
CF	11	7	12	51	81
PF	0	4	0	15	19
MD	0	6	26	82	114
SF	0	1	0	2	3
NO	0	0	5	14	19
Column Totals	11	18	43	164	236

ASSEMBLAGE # 15A -- STAGE II BIFACE REDUCTION (hard hammer)

	XL	L	M	S	Row Totals
CF	0	3	4	15	22
PF	0	2	0	6	8
MD	0	5	17	22	44
SF	0	0	1	0	1
NO	0	0	1	0	1
Column Totals	0	10	23	43	76

ASSEMBLAGE # 21A -- STAGE II BIFACE REDUCTION (hard hammer)

	XL	L	M	S	Row Totals
CF	0	7	22	20	49
PF	0	5	3	4	12
MD	1	4	22	64	91
SF	0	0	0	4	4
NO	0	0	4	5	9
Column Totals	1	16	51	97	165

ASSEMBLAGE # 21B -- STAGE III BIFACE REDUCTION (soft hammer)

	XL	L	M	S	Row Totals
CF	2	6	22	94	124
PF	0	3	11	34	48
MD	0	4	38	154	196
SF	0	0	3	4	7
NO	0	0	4	22	26
Column Totals	2	13	78	308	401

ASSEMBLAGE # 16 -- STAGE III BIFACE REDUCTION (soft hammer)

	XL	**L**	**M**	**S**	**Row Totals**
CF	0	2	15	50	67
PF	0	3	12	46	61
MD	0	4	40	124	168
SF	0	0	4	16	20
NO	0	0	9	44	53
Column Totals	0	9	80	280	369

FLAKE COUNTS
according to Magne's Scar Count Approach

Legend:
N= total # of flakes in the assemblage

E= "Early" stage debitage (i.e.,0-1 flake scars)
M= "Middle" stage debitage (i.e., 2 flake scars)
L= "Late" stage debitage (i.e., 3 or more flake scars)

PRB= Platform Remnant Bearing flake
BRF= Biface Reduction Flake
BPO= Bipolar flake

OBSIDIAN EXPERIMENTS

	BIPOLAR	BIPOLAR	MED. CORE	MED. CORE	LG. CORE
ASSEM. #	13	14	11	19	17
N	142	92	136	111	142
E-PRB	3	1	26	12	18
M-PRB	0	0	0	9	1
L-PRB	1	0	3	2	0
E-SHATTER	57	19	60	58	57
M-SHATTER	51	22	31	29	43
L-SHATTER	26	32	14	7	24
BRF	0	0	1	0	0
BPO	0	18	0	0	0

	LG. CORE	STAGE 2	STAGE 2	STAGE 3	STAGE 3
ASSEM. #	18	15A	21A	16	21B
N	209	100	161	336	309
E-PRB	18	12	27	16	17
M-PRB	14	7	15	28	11
L-PRB	48	5	10	31	13
E-SHATTER	64	25	37	80	60
M-SHATTER	39	31	47	77	78
L-SHATTER	26	20	22	89	104
BRF	0	0	0	15	17
BPO	0	0	0	0	7

FLAKE COUNTS
according to Magne's Scar Count Approach

Legend:
N= total # of flakes in the assemblage

E= "Early" stage debitage (i.e.,0-1 flake scars)
M= "Middle" stage debitage (i.e., 2 flake scars)
L= "Late" stage debitage (i.e., 3 or more flake scars)

PRB= Platform Remnant Bearing flake
BRF= Biface Reduction Flake
BPO= Bipolar flake

QUARTZITE EXPERIMENTS

	BIPOLAR	BIPOLAR	MED. CORE	MED. CORE	LG. CORE
ASSEM. #	12	2	3	20	5
N	111	90	91	124	219
E-PRB	1	8	26	22	28
M-PRB	1	1	0	1	14
L-PRB	0	1	0	0	5
E-SHATTER	69	50	49	53	101
M-SHATTER	27	17	7	34	55
L-SHATTER	4	9	9	13	13
BRF	0	0	0	1	2
BPO	8	3	0	0	0

	LG. CORE	STAGE 2	STAGE 2	STAGE 3	STAGE 3
ASSEM.#	7	6A	8A	6B	8B
N	192	118	165	71	107
E-PRB	38	18	23	16	28
M-PRB	0	7	0	7	4
L-PRB	2	0	1	1	8
E-SHATTER	91	62	100	29	27
M-SHATTER	44	26	35	15	21
L-SHATTER	17	5	4	2	12
BRF	0	0	0	1	7
BPO	0	0	0	0	0

FLAKE COUNTS
according to Ahler's Mass Analysis

Legend:
G1= size grade 1 (1" screen)
G2= size grade 2 (1/2" screen)
G3= size grade 3 (1/4" screen)
G4= size grade 4 (1/8" screen)

N= total number of flakes/assemblage

REDTYPE= type of reduction experiment

OBSIDIAN EXPERIMENTS

REDTYPE	N	G1 COUNT	G2 COUNT	G3 COUNT	G4 COUNT
Bipolar/#13	500	8	32	101	359
Bipolar/#14	336	3	21	68	244
MedCore/#11	448	10	25	101	312
MedCore/#19	803	16	23	72	692
Lg.Core/#17	635	23	19	100	493
Lg.Core/#18	561	18	27	164	352
Stage 2/#15a	463	1	28	73	361
Stage 2/#21a	472	13	34	114	311
Stage 3/#16	1539	6	63	266	1204
Stage 3/#21b	1455	12	57	262	1074

Total # Flakes: 7290

QUARTZITE EXPERIMENTS

REDTYPE	N	G1 COUNT	G2 COUNT	G3 COUNT	G4 COUNT
Bipolar/#2	221	8	18	64	128
Bipolar/#12	401	5	24	82	288
MedCore/#3	105	4	29	54	15
MedCore/#20	418	8	30	86	288
Lg.Core/#5	724	11	53	154	501
Lg.Core/#7	653	22	46	127	456
Stage 2/#6a	391	10	76	32	270
Stage 2/#8a	488	14	56	95	319
Stage 3/#6b	187	2	28	41	116
Stage 3/#8b	700	11	53	148	486

Total # Flakes: 5152