ARCHAEOLOGICAL PETROLOGY AND THE ARCHAEOMETRY OF LITHIC MATERIALS*

M. S. SHACKLEY

Geoarchaeological XRF Laboratory, Department of Anthropology, University of California, Berkeley, CA 94720, USA

For 50 years, archaeologists and physical scientists have been dating, determining the composition of and measuring stone tools, and reporting them in Archaeometry and many other journals. In Archaeometry specifically, the number of papers devoted to the analysis of lithic material has increased at least 30 times since 1958 and volume 1. This is a reflection not only of an increase in the number of scholars devoting their time to the archaeometry of stone, but also of increases in the quality and quantity of instrumental technology available to researchers in the field.

KEYWORDS: ARCHAEOLOGICAL PETROLOGY, LITHICS, OBSIDIAN, CHERT, MARBLE, INSTRUMENTAL TECHNIQUES, SAMPLING

INTRODUCTION

In 1983, D. R. C. Kempe and Anthony Harvey published an edited volume entitled The petrology of archaeological artefacts. This was the first attempt to compile the ever-increasing work on the geoarchaeological analysis of stone in all its forms from chipped to sculptural stone, and it included chapters on ceramics and metals. The latter two data sets are not part of this paper; they are covered very well by others in this set of special review papers. This was a landmark study that received very little attention by lithic technologists, particularly in the Americas.

The purpose here in the space allotted is to examine a bit of the history and advances made in the archaeometry of stone, with special emphasis on the direction that a 21st-century archaeometry is currently making towards an understanding of the composition of stone artefacts.

From chipped stone to sculptural stone, by both optical methods (petrography) and analytical chemistry, the quantity and quality of research has increased exponentially. This is, to a limited extent, an expanded view of a similar review about 10 years ago (Shackley 1998a). There is no attempt to cover dating stone, or lithic technology, in this case defined as the metric and technological analysis of stone as an artefact produced by humans through time (debitage analysis, use-wear studies etc.), except where archaeometric techniques are beginning to shed light on technological issues (cf., Grace 1996). This is a huge literature, and beyond the focus of archaeometry in my estimation (for chert, see Delage 2003). However, as I will note throughout, lithic technologists have much to gain from those of us who are focused on the archaeometry of stone. Many lithic technologists have learned this in the past decade or so, and all of us are better for it (Shackley 1998a; Odell 2003). Unfortunately, most lithic technologists, particularly those in the Americas, have not even heard of Kempe and Harvey’s (1983) volume. To this, as a geoarchaeological scientist and a lithic technologist, I devote this paper.
Disclaimer

As with many of the subjects covered in these relatively short summaries, the entirety of archaeological petrology and stone is impossible to cover. The goal here is to address some of what I consider to be the important recent advances, as well as some of the continuing problems that affect archaeological geochemical studies of stone, with the archaeologist in mind rather than the trained geochemist. There are a number of important new and experimental techniques in archaeometry, both in the field and laboratory, that are being applied to stone tools and structural stone studies in archaeology, and some will wonder why their favourite technique or perspective was not covered (such as the exciting new possibilities of isotope studies through ICP–MS; see Speakman and Neff 2005). I want to emphasize that this is my particular view, and not necessarily that of other archaeologists or archaeometrists. I find the new techniques and subject areas discussed here to be of primary importance to the shifts in archaeological thought in the past decade, and the set of problems in chert, marble and obsidian characterization, one that simply will not and should not go away.

A SHORT HISTORY IN THE JOURNAL *ARCHAEOMETRY*

As suggested earlier, papers on analyses of lithics in *Archaeometry* are indicative of the trajectory of research on this subject in archaeology in general (see the Appendix and Fig. 1). The Appendix is a bibliography of all the 117 papers devoted to lithic material analysis from and including volume 1 (1958) through to volume 49 (November 2007). Again, this does not include those papers devoted solely to ‘lithic technology’ as I defined it above; there are very few of these. *The Journal of Archaeological Science* (JAS), published first in 1974, seems to accept more purely technological papers on lithics than *Archaeometry* has by comparison. As first an associate, then the Society of Archaeological Science Managing Editor at *Archaeometry* between 1999 and 2006, I would say that this is not due to editorial fervour or preference, but a perception by archaeologists that JAS is more accepting of technological studies.

Figure 1 exhibits the frequency of papers devoted to archaeometric analyses of lithic material, arbitrarily divided into decades. The earliest class is only from 1958 to 1959, and the latest class only runs from 2000 to 2007 for obvious reasons. The near exponential rise in this research is not due to editorial preference or the increase in page numbers (from two to four issues per year), but is a real reflection of the increases in lithic material research similarly reflected in other subject areas reported in this volume (see Delage 2003). However, it is noticeable that *Archaeometry* was dominated by ceramic and metals studies in the early days, still a prevalent ratio in many parts of the world and lamented by some (Williams-Thorpe 1995; Shackley 1998a, 2005).

THE CHEMICAL CHARACTERIZATION OF STONE: BEYOND FISHING EXPEDITIONS

*Define the problem. A good source analysis is usually complex and expensive, and should not be initiated merely as a fishing expedition in the vague hope that you might find something interesting.* (Luedtke 1992, 117)

Since a good part of my own research is involved in the wavelength and energy-dispersive X-ray fluorescence (EDXRF) analysis of obsidian artefacts, it seems odd that I would continually discuss what I call the ‘sourcing myth’ in archaeometry (Shackley 1998a, 2005). As many of
the archaeometrists who are involved in chemical characterization know, few archaeologists understand the physics and geochemistry of XRF, neutron activation analysis (NAA), proton-induced X-ray emission – proton-induced gamma ray emission spectrometry (PIXE–PIGME), mass spectrometry (MS), inductively coupled plasma – mass spectrometry (ICP–MS) or any of the other techniques used to characterize stone materials. Even more worrisome is the fact that few explore the issue of statistical probability and its role in assigning an artefact to source (see Baxter 1994; Neff and Glascock 1995; Shackley 1998b). This would all be fine if archaeometrists were all absolutely honest (and most are, most of the time), and there was a better line of communication between the scientist and the archaeologist—which there is not, or at least there has not been good communication. There are, of course, good reasons for this lack of communication. Due to heavy workloads and the sheer size of the printed world in archaeology, it is necessary that the archaeologist trusts the archaeometrist to be honest and that the source assignments are ‘right’.

In this review, the problems in the chemical characterization of three very important substances used throughout prehistory and the focus of much archaeometric research in the past three decades—marble, secondary siliceous sediments, called chert here, and volcanic glass, both silicic and mafic—are the focus. The issues raised are typical of the analysis of any stone material, particularly those that are heterogeneous.

One of the most misused terms in archaeometry is the word ‘sourcing’. Archaeologists most often use it, and unfortunately archaeometrists do too. Besides the grammatical problems with the word, it implies that whatever is submitted to the archaeometrist will return with a bona fide and certified source provenance that is not probabilistic at all, but confidently determined. I am certainly not the first to say this, but I will reiterate that nothing is ever really ‘sourced’. The best we can do is provide a chemical characterization and a probable fit to known source

Figure 1 The frequency of papers devoted to the archaeometric analysis of lithic material in Archaeometry in 10-year increments (the earliest and latest classes are truncated).
data. The key words here, of course, are characterization and known. So many regions have poor source standard data that we can never really be sure, even beyond the scientific process of orderly refutation of alternatives (Ward 1977; Luedtke 1992; Hughes and Smith 1993; Williams-Thorpe 1995; Cauvin and Balkan-Atlı 1996; Gratuze 1999; Kuzmin et al. 2002; Shackley 2005; Maniatis 2007).

Chert and characterization

Attempts to determine the source provenance of chert through chemical characterization have vexed archaeologists and archaeometrists for decades (Luedtke and Myers 1984; Matiskainen et al. 1989; Luedtke 1992; Warashina 1992; Church 1995; Hoard et al. 1995; Roll et al. 2005). This is mainly due to the processes of the formation of chert through precipitation from a parent material at ambient temperatures, where only certain elements and compounds can be removed from the original matrix. Additionally, sources of chert, particularly Phanerozoic marine cherts, by their nature often cover large geographical expanses, which decreases their utility as a means to derive inferences for exchange, interaction or procurement processes. Most of the techniques that have proven helpful in deriving useful chemical data for cherts are destructive and expensive, further limiting its utility, at least where conservation of objects is an issue (see Roll et al. 2005). Barbara Luedtke’s (1992) treatment of the subject for archaeologists is an excellent step towards generating an understanding of chert among archaeologists. Among the many useful ideas, Luedtke suggests a four-step process for determining the probable source of cherts (1992, Appendix A). I have modified these steps and added a few more based on my own experience, and indeed, in general, these steps are appropriate for most stone provenance studies:

1. A number of similar unknown sources may be detected from various archaeological contexts. Often, these will cluster in various geological regions.
2. Consult the regional geological literature to ascertain the location of the nearest and most likely rock bodies. In the case of obsidian, glass is rarely mentioned in association with mapped rhyolite bodies in the geological literature.
3. Frequently, local residents, archaeologists, rock hounds, rock hound guides and generally word-of-mouth have been found to be excellent sources for locating obsidian and chert. Nearly all of the obsidian and chert sources ‘discovered’ in the North American Southwest were first found by non-scientists (Shackley 2005).
4. Most often, it has been found necessary, after determining the general location of a source, to begin the search downstream from the probable source area. Many of the Tertiary and earlier sources are highly eroded, and many nodules and fragments of rock can be released into the sediment load and carried many kilometres downstream from the source (Shackley 1998c, 2005).
5. In the field, when the source or source area is discovered, the extent and density, general geological setting and variability of lithic production should be recorded. Pedestrian transects across the source area are sampled and form strata from which probability samples are selected for optical and/or instrumental analysis, ensuring as representative a sample as possible.
6. Determine which techniques or procedures are worth pursuing. For museum specimens, it is often unacceptable to thin section the material for petrographic analysis or subject it to destructive NAA, while non-destructive methods such as XRF may be acceptable. Keep in mind that each method has its advantages and limitations, and the results of a given technique may not be specifically valid (see Goffer 1980; Neff and Glascock 1995; Green 1998; Shackley 1998a; Roll et al. 2005; Speakman and Neff 2005).
(7) Finally, match artefacts to sources. The uses and abuses of performing this step are astronomical in the literature. Multivariate techniques are often sought after with the ease of PC-level software now available, but archaeologists often do not have the statistical background to evaluate these results. Starting with simple bivariate plots and central tendency statistics comparing the artefact and geological data is often sufficient to assign artefacts to sources (Hughes 1984a; Shackley 1988, 2005; Baxter 1992, 1994; Neff and Glascock 1995; Glascock et al. 1998).

(8) The source descriptions should include exact locations; preferably Universal Transverse Mercator international grid system (UTMs), a description of the geological context, megascopic attributes of the raw stone (cortical and interior variation, nodule size, colour, fabric and opacity), the distribution of secondary deposits, the density and character of lithic reduction at the source, and relevant published sources.

Due mainly to these issues, archaeometrists and archaeologists have in the past decade or so begun to focus on the very real problems of chert characterization and source assignment (Church 1995; Hoard et al. 1995; Roll et al. 2005). As archaeologists begin to gain an understanding of the instrumental techniques involved and the assumptions that archaeometrists often must make, they are beginning to require that archaeologists who employ archaeometric techniques for characterization pay heed to sampling design and presentation of results. Much of the wrangling apparent in two of the articles cited above revolved around issues of sampling at the source. Church rightfully questioned whether sampling reduction debris at the source was sufficient and necessary to deal with source variability. Hoard et al. noted that while some geological samples were gathered, much of the characterization project was based on second-hand collections. In this case the potential for committing Type I or Type II errors, mainly due to retooling at the source, can be great, and the archaeologist or the archaeometrist may not even be aware (Thomas 1986, 213–17; Bernard 2006). Without a firm knowledge of potential chert sources in a region, errors of misassignment will occur. Even with this knowledge, the level of macroscopic and chemical variability inherent in secondary siliceous sediments, confident characterization and source assignment can be hazardous. Given all these caveats, however, chert characterization and source assignment are just beginning to be a real tool in archaeology, using confocal microscopy tied to optical petrography, as well as the instrumental techniques mentioned here. While fraught with the difficulties of source heterogeneity and the general need to use destructive or partially destructive techniques, hope is in the air for the relationship between archaeometry and one of the most commonly used flaked stone raw materials in prehistory.

**Obsidian characterization and the ‘sourcing’ myth**

Many of the points briefly outlined for chert studies are equally relevant for obsidian studies in archaeology, but because these homogeneous, disordered, silicic glasses can be so precisely characterized, the potential abuses are much greater. Since it is a common aspect of archaeological practice and a good example for other data sets, a few words will be directed towards some of the problems in the chemical characterization of silicic glasses that directly impact the resolution of archaeological problems.

Obsidian studies in archaeology are a relatively recent aspect of archaeological research (Boyer and Robinson 1956; Green 1962, 1998; Cann and Renfrew 1964; Jack and Heizer 1968; Jack and Carmichael 1969; Williams-Thorpe 1995; Shackley 2005). Early on, megascopic (macroscopic) observation, density measures and mass spectrometry were all used...
in an attempt to define source groups and correlate artefacts to sources, with mixed results. In the late 1960s and early 1970s, with the evolution of relatively inexpensive X-ray fluorescence spectrometers and the increasing use of NAA, obsidian studies in archaeology began to be used in an effort to deal with issues of exchange and interaction (Glascock 1994; Shackley 1998b, 2005; Glascock et al. 2007). Archaeologists working in the Mediterranean, the New World and Oceania were particularly keen to use these new methods, mainly since the instruments were available at most university campuses (Hughes 1984b; Nelson 1984; Shackley 1988, 1995, 2005; Tykot 1992; Williams-Thorpe 1995; Green 1998). By the 1980s, archaeologists in every part of the world where obsidian occurred were engaged in obsidian provenance studies. Now, thousands—if not tens of thousands—of pieces of archaeological obsidian are analysed yearly by X-ray fluorescence spectrometry (XRF), neutron activation analysis (NAA or INAA), inductively coupled plasma – mass spectrometry (ICP–MS) and proton-induced X-ray emission – proton induced gamma ray emission (PIXE–PIGME), and much of the analysis is performed by archaeologists now rather than by physical scientists (Shackley 2002). This is the hope for the future, but there are limitations. As discussed above, while most archaeologists inherently know that a chemical signature is only as good as the instrument and, more importantly, the available source data, too few seem to worry about this. Many of the source assignment problems inherent in earlier semi-quantitative analyses, however, are simply gone now, with the emphasis on easily derived quantitative estimates.

The obsidian hydration controversy

Archaeometry has played a role in the issues surrounding obsidian hydration dating nearly since the beginning (Layton 1973; Ericson and Kimberlin 1977; Michels et al. 1983; Stevenson et al. 1987, 1989; Liritzis 2006). Obsidian hydration dating, however, has not played a major role in Old World chronology building for a number of reasons, and European scholars have not dealt with the controversy in a major way. For this and the controversy surrounding the method, I will not spend much time on the subject. It is most ‘popular’ in the New World, and mainly restricted to parts of North America (California, the Great Basin) and Oceania, particularly New Zealand. There are historical reasons for this—primarily, of course, a dominance of obsidian raw materials in those regions, and the absence of other datable materials. For many of us in the field, it has been fraught with too many issues to function as a reliable member of the chronometric toolkit (Shackley 2005).

Beginning with Friedman and Smith’s (1960) paper on dating using obsidian hydration, scholars have moved into two camps; those who believe fervently in the utility and validity of the method (Jackson 1984; Hall and Jackson 1989; Hull 2001, 2002) and those who find the recurring problems simply too much (Riddings 1996; Anovitz et al. 1999; Loyd 2002—cf., Stevenson et al. 1998). Even those, as above, who have found issues regarding the use of obsidian hydration as a direct dating method still work towards resolution of the issues. Factors such as variances in atmospheric heat and humidity through time, the inability to control the burial history of the artefact, resetting the hydration rim after post-depositional burning, seemingly exponential, non-linear shifts in the absorption of water over time and inter-observer measurement error all frustrate attempts to use hydration as an absolute method. More recently, Hull (2001) has argued that these critiques are simply because researchers focus on the method as a direct rather than relative technique, although Hull seems to use it as the former in her own work (Hull 2002). The recent symposium on the effects of fire on obsidian hydration in California found, through empirical observation and experimentation, that hydration
rims can be reset when subject to forest fire conditions, dealing a lethal blow, at least for the non-believers (Deal 2002; Loyd 2002; Loyd et al. 2002; Skinner 2002). Some western North American federal archaeologists will no longer pay for obsidian hydration analyses through federal funds (Skinner 2002, and pers. comm.). Parenthetically, through experimentation and in-field studies, it has been found that there is no statistically significant change in trace element chemistry when heated to high temperatures (Skinner et al. 1996; Shackley and Dillian 2002; Steffen 2002).

So, what should archaeologists do? Hull suggests that obsidian hydration should not be seen as an absolute dating method like dendrochronology, but as a relative dating tool for application in, for example, site contexts where the depositional history is similar from one area to another (Hull 2001, 1026). In recent Historic-period sites, Hull has been able to derive some interesting relative dates in a chronological context where $^{14}$C dating is useless, given the inherent uncertainty of the calibration curve in near ‘modern’ periods (Hull 2002; see also Ambrose 1998). Even so, these dates have yielded error rates due to the vagaries of hydration. Riddings (1996) and Anovitz et al. (1999) seem to be willing to claim obsidian hydration a failure. I am glad to see, however, that some, including Anovitz, are willing to continue experimentation in the hope that the issues will be resolved (Friedman et al. 1997; Stevenson et al. 1998).

Fun with marbles

Aside from obsidian provenance studies, marble is one of the success stories in lithic provenance work in archaeology, or at least scholars are intensively focused on resolving the issues (Herz and Waalkens 1988; Polikreti and Maniatis 2002; Luke et al. 2006; Maniatis 2007). Archaeologically, marble is one of the most important metamorphic rocks—if not the most important. Fine-grained varieties have been used for sculptural and architectural stone throughout history, in both the Old and New Worlds (Garrison 2003). In the classical Greco-Roman tradition, marble became the preferred building stone, and it continues to be an important interior and exterior stone.

Marble, as a granular, heterogeneous rock primarily composed of calcite or dolomite, derived from limestone, is subject to many of the same heterogeneity issues as chert. The colour and texture, which are sometimes quite variable, are the result of precursor minerals and metamorphism. The lower the concentration of accessory minerals, the whiter is the marble (Maniatis 2004). As in many early studies, NAA was seen as the saviour of marble analysis, and would simply solve all the issues. However, as discussed above, NAA, like XRF, is a mass analysis technique, and all elements for all the constituents are included in the analysis. Again, as in chert studies, it is now accepted that multiple analyses are necessary to confidently characterize marbles (Polikreti and Maniatis 2002; Maniatis 2004; Luke et al. 2006). Other techniques that are frequently used—or at least were used in the past—include: stable isotope analysis by mass spectrometry, focused on $^{13}$C/$^{12}$C and $^{18}$O/$^{16}$O ratios; cathodoluminescence (CL); optical petrography; and most recently electron paramagnetic resonance spectroscopy (EPR), also known as ESR. Polikreti and Maniatis (2002) note that EPR is very effective at detecting manganese (Mn$^{2+}$) and iron (Fe$^{3+}$) ions, which exist in various amounts in every marble crystal, replacing calcium (Ca$^{2+}$). The proportions of these ions follow the conditions present during metamorphism, and therefore are very locally source specific (see also Maniatis 2007). This resolves some of the stable isotope difficulties where source overlap problems are not solved. While EPR has been shown to be quite successful in separating many marble sources in the Old World, it still relies on an extensive source database, and does not solve all
source overlap. However, it does look as though—in the Old World at least—a comprehensive
and usable database for marble is nearly here. If chert were not so widespread, as marble, this
would be the case for that stone too.

**Sampling: how much is enough?**

Sampling issues in archaeology are, of course, not new. In the 1980s, the US Department of
the Interior spent a considerable sum investigating the issue, both for confidence in predictive
modelling and for regional and intra-site sampling (Judge and Sebastian 1988). Running
throughout this work, and many others, and a basis for probability theory, is the necessity for
adherence to the simple law of large numbers: as the sample size increases, the probability of
attaining a representative sample increases, diversity issues aside. Given the funding con-
straints on archaeology in all forms, the law of large numbers is an ever-present demon. The
sampling issue deserves some thought at all levels of archaeometric analysis, both in the field
and the laboratory, and while we may not always be able to fund projects based on these
precepts, we at least need to be familiar with the risks inherent in our decisions. In both source
sampling and sampling artefacts for analysis from an archaeological assemblage, there appears
to be a minimum number necessary to derive a confident conclusion. Again, a couple of examples
at both the source and artefact assemblage level from my own work in the North American
Southwest are offered.

**The Mule Creek regional source: ‘real’ chemical source variability**

In a 1988 study of the Mule Creek source in western New Mexico, sampling 15 source standards
from a population of 200 gathered at five localities, two chemical ‘outliers’ were noted,
collected with significantly higher rubidium concentration values (Shackley 1988a, 767).
These outliers have now been identified as a distinct chemical group, often mixed in the
regional Gila Conglomerate with three other chemical groups (see Shackley 1995, 2005). The
eruptive geology in the area is complex, and has been studied by Ratté and others for some
time (for a summary of this work, see Ratté 2004). Primary *in situ* perlite localities with remnant
marekanites for three of the chemical groups have thus far been located.

At least four distinct chemical groups are evident, distinguished by the incompatible
elements Rb, Y, Nb and Ba, and to a lesser extent Sr and Zr, and are named after the localities
where marekanites have been found in perlitic lava: Antelope Creek, Mule Mountains and
Mule Creek/North Sawmill Creek, all in New Mexico. Additionally, during the 1994 field
season, a fourth subgroup was discovered in the San Francisco River alluvium near Clifton,
Arizona, and in older alluvium between Highway 191 and Eagle Creek in western Arizona,
north of Clifton. While *in situ* nodules have not yet been found, they are certainly located
somewhere west of Blue River and north and west of the San Francisco River, since none of
this ‘low zirconium’ subgroup was discovered in alluvium upstream from the juncture of the
Blue and San Francisco Rivers. The genetic relationship is apparent in the bivariate and trivariate
data plots (Shackley 2005, 51, figure 3.6), and signifies the very complex nature of the Mule
Creek silicic geology, with subsequent depositional mixing in the Gila Conglomerate. The
original simple random sample of 15 nodules merely indicated that some variability existed,
and only after extensive transect survey and sampling, and locating three localities with *in situ*
marekanites did it become apparent that there was significant chemical variability inherent in
the ‘source’. Perhaps most importantly, without the knowledge of these other subsources,
analysis of archaeological obsidian in the region could posit a new and unknown source. This process has been repeated throughout the world, and has really changed the way in which we view sources of archaeological stone (Hughes and Smith 1993; Williams-Thorpe 1995; Cauvin et al. 1998; Glascock et al. 1998; Summerhayes et al. 1998; Tykot 1998; Gratuze et al. 2005; Poupeau et al. 2005; Shackley 2005).

So, how much should be sampled? As usual, the answer depends on many factors. Often, the archaeometrist has detected variability in the source from archaeological analyses even though source standards have not indicated that variability (Hughes and Smith 1993; Glascock et al. 1998; Kuzmin et al. 2002; Gratuze et al. 2005; Poupeau et al. 2005; Carter et al. 2006; Carter and Shackley 2007). No longer are 10 samples sent to an archaeometrist sufficient to confidently define the elemental variability of a source. The following source sampling strategy is a guide [composed of our own work in the American Southwest; the work of Glascock and others (Glascock et al. 1998) in Mesoamerica; Skinner’s (1983) work in the American Northwest; Jackson’s (1986, 1989) and Hughes’ (1988, 1994; see also Hughes and Smith 1993) work in California and the Great Basin; the work of Torrence and others (Torrence 1992; see also Summerhayes et al. 1998) and Green (1998) in Oceania; and the work of Carter et al. (2006), Gratuze et al. (2005), Poupeau et al. (2005), Williams-Thorpe (1995) and Tykot (1998) in the Mediterranean and the Near East]:

1. On the ground, sample surveys must include the entire primary and secondary extent of the source. This often requires days or weeks of field reconnaissance, and extensive discussions with geologists and the local populace.

2. Often, in order to determine the probable location of primary sources, extensive sampling of the secondary distribution can isolate probable areas for further study—the San Francisco River Alluvial samples were delimited in this way (see above).

3. All samples should be located at least to the section, square kilometre, UTM or global positioning system (GPS) coordinates, depending on the country of origin.

4. Stratified subsamples should be derived in the laboratory by the sublocalities collected in the field (preferably analysing at least five or more samples from each locality—but note that this is not a magic number). This could mean taking thousands of samples in the field.

The above is only an outline for research. In some instances, particularly with Quaternary sources, secondary deposition may not be an issue, since there has been little geological time for the glass to erode into the environment and/or large-scale pyroclastic eruptive events may not have occurred (see Shackley 2005). While this eliminates the problems of secondary deposition, it does not allay the problems of possible intra-source chemical variability. While this strategy is focused on obsidian sources, most of the steps are equally valid for any other stone material, including marble, other volcanics and secondary siliceous sediments.

INSTRUMENTATION: WHICH INSTRUMENT IS BEST?

Just about the most frequently asked question by archaeology students is: ‘Which instrument is best to analyse my stone objects?’ The answer, unfortunately, is: ‘It depends . . .’ Again, the problem design and the level of precision needed to address that design will determine which instrument is the best for a given project. Ten years ago, this question was a bit easier to answer (Shackley 1998b).

With the general improvements in technology, instrumental techniques in archaeological geochemistry have similarly improved. Almost all instrumental and empirical techniques have been used, including density, magnetism, atomic absorption, PIXE–PIGME, ICP (for obsidian,
see Tykot and Young 1995; Shackley 2005; Speakman and Neff 2005), megascopic criteria and others. Today, three major instrumental methods dominate the field: neutron activation analysis (NAA or INAA), wavelength- and energy-dispersive X-ray fluorescence (XRF), including portable XRF (PXRF), and increasingly LA–ICP–MS (Neff and Glascock 1995; Glascock et al. 2005; Shackley 2005; Speakman and Neff 2005). All of these methods have benefited from the revolution in microprocessors and the attendant software industry, and are easier to use, and hence misuse, today. In this summary, there is not the space to detail the intricacies of each method. For archaeologists, Goffer (1980) and, to a certain extent, Harbottle (1982) do this quite well for the earlier period. Glascock (1991) presents a good, though technical, treatment of NAA (see also Neff and Glascock 1995, and see Speakman and Glascock’s 2007 special issue on NAA in Archaeometry), and XRF and PIXE–PIGME are explained in some detail in Shackley (1998b) (see also http://www.swxrflab.net/xrfinstrument.htm). The goal here is to address the relative merits of the three most commonly used analytical techniques, NAA, EDXRF and PIXE–PIGME, given the very real situations encountered by archaeologists. Marble analyses have gone through even more experimentation with regard to instrumentation, as briefly discussed above.

There has been a certain level of mythology regarding the optimal analytical instrument for geochemical studies of stone. The prevailing ideal seems to be that NAA is the ‘best’ way to go if funding is not a problem. NAA for most elements is certainly precise and it simply can detect more elements than the other two methods (Glascock 1991; Neff and Glascock 1995; Speakman and Glascock 2007). Neutron activation analysis, however, has two primary short-comings. It is in essence a destructive technique, and while the material is not technically destroyed, depending on original size, it may be broken into relatively small pieces and remain radioactive for many years. Additionally, partly due to cost and partly due to the public fear of nuclear energy, NAA is not readily available. Also, for obsidian characterization, NAA cannot analyse for Ba, Sr and Zr as accurately as the other two methods (see Shackley 2005; Craig 2007). These three elements are important incompatibles in silicic melts and can be extremely important in separating sources or dealing with intra-source issues (Hildreth 1981; Mahood and Hildreth 1983; Macdonald et al. 1987; Hughes and Smith 1993). Neutron activation can, however, analyse for a number of other incompatibles quite accurately that are outside the range of XRF.

For museum specimens and artefacts that are subject to repatriation, NAA is not a feasible choice. If precision and accuracy are necessary, which can be an issue in intra-source studies and where the possibility of long-distance exchange is probable, then NAA will always provide the most efficient alternative. Additionally, if the sample is extremely small (< 5 mm in largest diameter), NAA or LA–ICP–MS may be the best analytical method. Other methods may require a minimum sample size for optimal results (Davis et al. 1998).

For a number of years now, Michael Glascock at the University of Missouri Research Reactor Facility, using NAA, and I at the Berkeley EDXRF Laboratory have been analysing the same obsidian and source standards and artefacts simply as part of our routine data sharing. It is remarkable how similar the results of the two analyses are, for the 17 or more elements measured in common (Shackley 1998a). In his paper on the chemical variability in the Coso Volcanic Field, Hughes (1988) also obtained similar results in the mid-Z incompatible elements between NAA and EDXRF. Much of this is due to the extreme accuracy of EDXRF in the mid-Z (and part of the high-Z) region, which fortunately contains some of the most sensitive incompatible elements for discriminating volcanic rocks, including Rb, Sr, Y, Zr, Nb and Ba (Giaque et al. 1993). Much of this ‘paired accuracy’ is due to calibration to international standards and the increasing analytical precision of today’s instruments.
What this all means for an archaeologist is that any laboratory employing EDXRF, WDXRF, NAA, PIXE–PIGME or ICP–MS will provide valid and comparable results, given a sufficient sample size, particularly for obsidian. The particular project constraints are now the main criteria for the selection of analytical methods. What is probably more important now is the accuracy of the source standard data, given the caveats throughout this paper. If the region of interest has not received the level of geoprospection necessary for accurate source assignment, it really does not matter how precise the instrument employed is known to be.

ARCHAEOMETRY AND THE INFERENCE OF TOOL USE

For most of its intellectual history, the understanding of prehistoric lithic technology has been dominated by direct observation, experiment and, fortunately or unfortunately, opinion; and just as dominated by polarized dialectic as to which particular paradigm was correct, or even ‘politically correct’ (Grace 1996). Many of these arguments have been based on opinions about the method and the validity of the assumed analogy between modern experiments and prehistoric knappers and tool users (Keeley 1980; Odell and Odell-Vereeken 1980; Gendel and Pirnay 1982; Unrath et al. 1986; Lewenstein 1987; Bamforth et al. 1990; Young and Bamforth 1990). The problems of inferring use from marginal damage on stone tools, particularly utilized flakes, has produced a rather high level of consternation among lithic technologists, and paramount frustration among contract archaeologists who are most pressed for time in concluding site function, and therefore site significance, in management decisions (see Grace’s 1996 review of the literature in Archaeometry).

One of the most commonly practised techniques, and one of the most empirically frustrating, is inferring use-wear and therefore function on unmodified flakes. Studies that rely on readily available microscopic techniques are exceptions rather than the rule. Most archaeologists rely on macroscopic examination of flake margins, deriving inferences concerning use on the basis of these gross morphological indications. Frighteningly, blind tests of various researchers examining experimental specimens macroscopically have indicated that few archaeologists agree on the character of the damage and hence the implied function of the tool (Bamforth et al. 1990; Young and Bamforth 1990; Christensen and Walter 1992; Christensen et al. 1992).

It is true that the more ‘experienced’ observers more often correctly identified the use-wear; some did not. This is considerably disturbing, since the most common flaked lithic tool found in prehistoric contexts is often the utilized flake. Young and Bamforth argue that ‘some degree of experimental experience should be required of all archaeologists who intend to include such tools in their research, simply to ensure that they observe their collections accurately’ (1990, 408). Gaining some experience with archaeometric methods should carry equal importance (see Burroni et al. 2002). Young and Bamforth may be over-optimistic and the error rate using macroscopic criteria is probably too high.

So what can archaeometry and analytical chemistry offer to solve the dilemma? Some studies by Marianne Cristensen, Phillipe Walter and others in France using environmental scanning electron microscopy (ESEM) may have far-reaching implications for use-wear research, and more recent studies using a variety of techniques have refined their effort (Christensen et al. 1998; Evans and Donahue 2005). Generally, even if you can determine the scraping, cutting, rotary-wear utilization with some degree of confidence, it is rarely possible to determine the actual material that the tool was used on beyond ‘hard or soft material’. Bone and wood are often lumped together under one material, but it is obvious that there are often important behavioural differences between stone tool use on either of these two materials.
Christensen et al. (1992—see also Christensen and Walter 1992; Menu and Walter 1992) have focused on margin polish on various stone tools by replicating tools and employing ESEM with an EDXRF probe to yield qualitative elemental data that can be paired with various substances. The results were cross-checked using RBS (Rutherford backscattering spectrometry) and PIXE (proton-induced X-ray emission spectrometry). Another aspect of their research that will not be treated in depth here is the empirical evidence for ‘melting silica’ where the polish is essentially a combination of loss of lepispheric quartz from the chalcedony matrix in chert and an accumulation of material, in this case bone (Fig. 1; Christensen et al. 1992, 491; see also Witthof 1967).

What is apparently important from this work is that the material that the tool was actually used on can be empirically defined. Quite similar and reliable data were also obtained with ESEM–EDXRF. Christensen et al.’s (1992) pilot study was applied to 13 000-year-old Magdelenian burins recovered from the important Paris Basin site of Etiolles, which was thought to be a lithic production area. Their study indicated that some of the burins were used to modify ‘chalky matter’, suggesting that they used the burins to remove the chalk cortex before early stages of reduction (Christensen et al. 1992, 492–3). Perhaps more illuminating was the analysis of Pre-dynastic Egyptian fishtail knives and ripple-flaked knives, which have often thought to have been used only for ritual purposes and not household or other utilitarian functions. Christensen et al. concluded, using this technique, that the fishtail knife was used to cut meat and the ripple-flaked knife was used in plant cutting (1992; Menu pers. comm. 1993). Most importantly, these tools were museum pieces that were thousands of years old, and that had been handled by many different people throughout their over one hundred years of museum life, and the technique is completely non-destructive when using the ESEM. This technique, while still experimental, is readily available to most archaeologists worldwide, and promises to solve some important issues that macroscopic and low-power microscopic techniques cannot.

This study has been replicated and refined more recently using LA–ICP–MS by Evans and Donahue (2005). The authors not only used LA–ICP–MS to great effect, but used even harsher cleaning techniques than the Christensen group a decade earlier and found that there was still not significant removal of the organic material (cf., Burroni et al. 2002). While ESEM and RBS are useful in determining the material targeted by the tool users in prehistory, LA–ICP–MS is more efficient, and given that the instrumentation is becoming readily available at many universities, promises to revolutionize lithic use-wear studies in the 21st century.

**AND NOW FOR THE FUTURE . . .**

There is always a risk of hoisting oneself on one’s own petard when forecasting the future, but there are a few comments that can be made with regard to the future of archaeological petrology and lithic studies. First, it is apparent that we are on the right track. Few people think that a single analytical method is sufficient to characterize any stone material—this is a good thing. Even obsidian, which is by definition completely disordered, with no crystalline structure, has been shown to often exhibit source heterogeneity. While XRF is a good non-destructive tool for characterizing obsidian sources and artefacts, it may not always include enough elements to solve all the problems. There are many regions in the world (northern Mexico, Eurasia, and even parts of the United States and Canada) where ‘unknown’ sources are still evident in the archaeological record.

Chert is another story. Many chert ‘sources’, particularly the Phanerozoic marine sediments, cover very large regions (i.e., the Ouachita Transition of the Edwards Plateau of Texas, and the
region of central France), and the consequent elemental and isotopic variability is so great that it may be a long time until chert characterization studies approach those of obsidian or even marble. As for marble, given that the sources are much more localized than for chert, and archaeometrists have been working with marble for decades now, the future looks very bright (Y. Maniatis pers. comm. 2007).

There is one point that bears repeating: megascopic characterization and source assignment on any rock is hazardous at best, and in many cases simply does not work (Shackley 2005, 101–5). We are long past the need for megascopic source characterization, and good riddance. Many of the techniques used to characterize rocks that have been discussed most here are relatively inexpensive, and in most cases there is no excuse to resort to megascopic speculation.

The Old versus the New World

I have not really spent much time discussing the very real differences between archaeometric practice and pedagogy in the Old World versus the New with regard to lithic material. I have talked to many on both sides of the Atlantic about this over the years, and especially Dave Killick at the University of Arizona, who is British but gained his Ph.D. at an American university. Now in press, Killick has a paper (2008) that every archaeologist on both sides of the Atlantic should read. In the United States and Canada, there are virtually no dedicated archaeological science programmes at any public or private university. There are those of us, such as Dave and I, who teach archaeological science as undergraduate and graduate method courses and have developed dedicated laboratories, but Americanist archaeology has fallen well behind Europe in dedicating courses of instruction and conferring degrees in the field (Calgary does have a graduate programme in geoarchaeology). Just as Europe cannot rely on the United States to supply brainpower, and should not, the lack of archaeological science programmes is a weakness of North American anthropological archaeology programmes, which seemingly rely on Europe to lead in this area. There are a number of social and structural reasons for this lack in US institutions; dominance of social anthropologists in anthropology departments, a continuing lack of rigorous science instruction in US K-12 schools; the religious right’s and postmodern philosophy’s attacks on science in the US; and very real funding issues. There really is not space to deal with these points here (see Killick 2008). These are some of the reasons why I suggest to my undergraduate students who are interested in archaeological science that they should look towards Europe for their graduate schools. The unfortunate outcome, however, for these students, is that US universities will not be interested in hiring archaeological scientists in anthropology departments. Still, many of us carry on, and very much look to the Old World as far as models for archaeological science programmes are concerned.

At 50 years old, *Archaeometry* is still the premier journal for archaeometry and archaeological science. With the proposed shift to nearly monthly publication of *Archaeometry*, the completion of digitization of all articles since 1958 and the journal’s great editing staff, lithic material studies and archaeological petrology in archaeology are in good hands. We still have many unanswered questions in the lithic realm, but I am confident that in time all of them will be addressed.

ACKNOWLEDGEMENTS

Reviews like this are not the work of one scholar. I benefitted immensely from advice on marbles and stone archaeometry generally from Yannis Maniatis. Obsidian, chert and marble geochemistry
have suffered through many of the same obstacles, and I value our discussions about stone and archaeometry. The undergraduate and graduate students in my lithic technology, archaeological science and archaeological petrology summer field school classes have taught me as much about geoarchaeological science as I have ever learned on my own. I want to thank Mike Tite for having the foresight to make the first grand changes at *Archaeometry*, instituting Managing Editors from the major archaeological science organizations worldwide, and Mark Pollard for continuing that trend—all the issues of *Archaeometry* are now digitized. *Archaeometry* has joined the 21st century!

**REFERENCES**


Green R. C., 1962, Obsidian: its application to archaeology, New Zealand Archaeological Society Newsletter, 5, 8–16.


Archaeological petrology and the archaeometry of lithic materials


APPENDIX

Archaeological petrology/lithic (stone) papers in Archaeometry, 1958–2007

This compilation of 117 papers that have appeared in Archaeometry over the past 50 years concerns the analysis of stone, but does not include any papers that cover purely lithic technology, and there are not many that cover the technology of stone tools. The list includes all stone analyses from the first issue in 1958 through to November 2007. The citations do not
include the journal title *Archaeometry*. If the stone material is not mentioned in the title, it is noted in square brackets.


Antonelli, F., Bernardini, F., Capedri, S., Lazzarini, L., and Montagnari Kokelj, E., 2004, Archaeometric study of protohistoric grinding tools of volcanic rocks found in the Karst (Italy–Slovenia) and Istria (Croatia), *46*, 537–52.


Della Casa, Ph., 2005, Lithic resources in the early prehistory of the Alps, 47, 221–34.


Hall, E. T., 1960, X-ray fluorescent analysis applied to archaeology, 3, 29–37 [general technique].


Holmes, L. L., and Harbottle, G., 2003, In the steps of William the Conqueror: neutron activation analysis of Caen stone, 45, 199–220.


Jones, M. C., and Williams-Thorpe, O., 2001, A illustration of the use of an atypicality index in provenancing British stone axes, 43, 1–18.


Knutson, K., and Hope, R., 1984, The application of acetate peels in lithic use wear analysis, 26, 49–61.


Lazzarini, L., and Antonelli, F., 2003, Petrographic and isotopic characterization of the marble of the Island of Tinos (Greece), 45, 541–52.


Mandal, S., 1997, Striking a balance: the roles of petrography and geochemistry in stone axe studies in Ireland, 39, 289–308.


Negash, Agazi, and Shackley, M. S., 2006, Geochemical provenance of obsidian artefacts from the MSA site of Porc Epic, Ethiopia, 48, 1–12.


Newman, R., 1992, Applications of petrography and electron microprobe analysis to the study of Indian stone sculpture, 34, 163–74.


215


Santi, P., Antonelli, F., and Renzulli, A., 2005, Provenance of medieval pieta ollare artefacts found in archaeological sites of central-eastern Italy: insights into the Alpine soapstone trade, 47, 253–64.


Uchida, E., Cunin, O., Shimoda, I., Suda, C., and Nakagawa, T., 2003, The construction process of the Angkor monuments elucidated by the magnetic susceptibility of sandstone, 45, 221–32.


Ward, G. K., 1974, A systematic approach to the definition of sources of raw material, 16, 41–54 [obsidian sources].


Williams-Thorpe, O., and Thorpe, R. S., 1990, Millstone provenancing used in tracing the route of a fourth-century BC Greek merchant ship, 32, 115–38.


Williams-Thorpe, O., Jones, M. C., Tindle, A. B., and Thorpe, R. S., 1996, Magnetic susceptibility variations at Mons Claudianus and in Roman columns: a method of provenancing to within a single quarry, 38, 15–42.