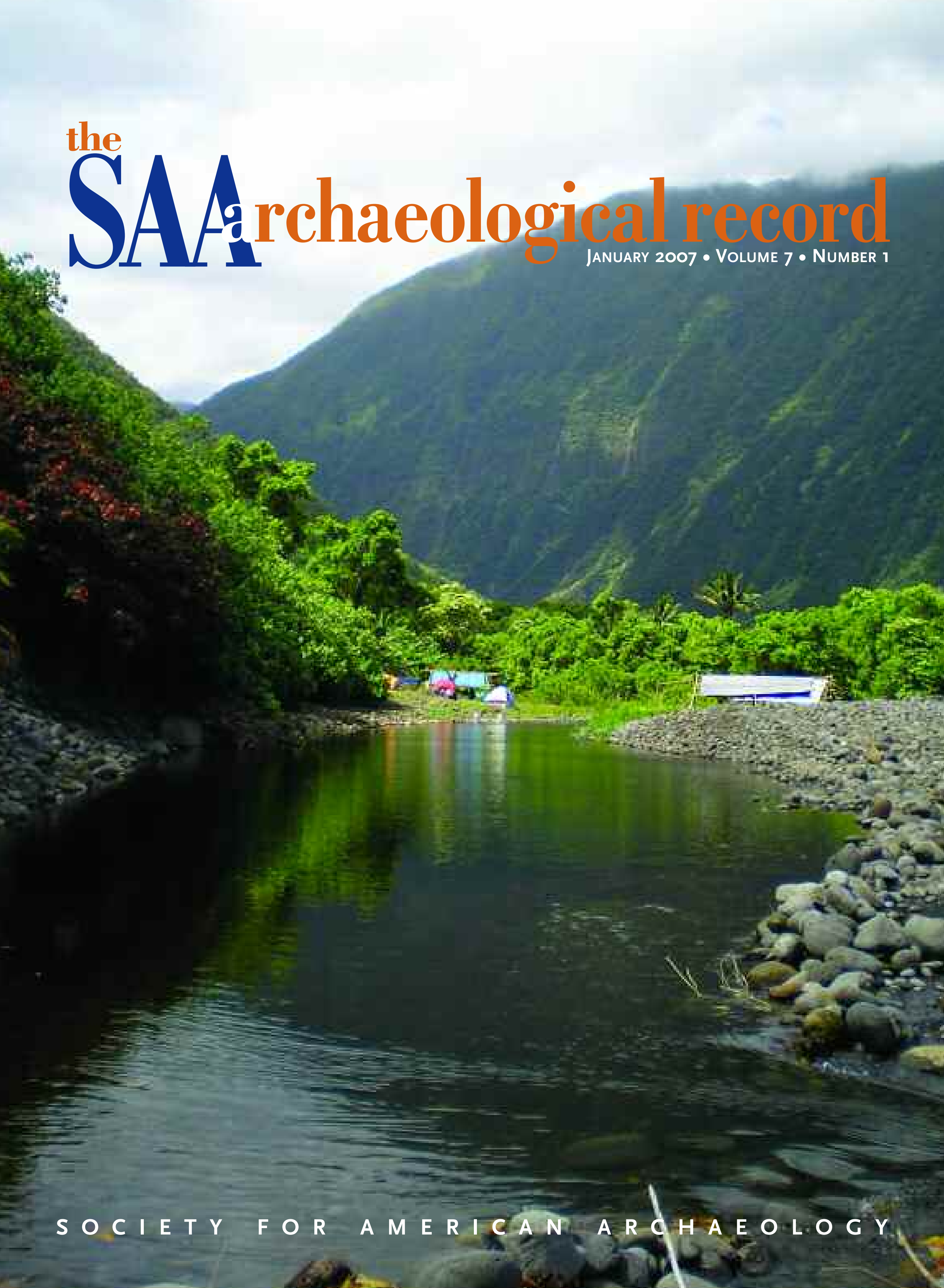
the SAA archaeological record

January 2007 • Volume 7 • Number 1



SOCIETY FOR AMERICAN ARCHAEOLOGY

THE USE OF BARCODES IN EXCAVATION PROJECTS

EXAMPLES FROM MOSSEL BAY (SOUTH AFRICA) AND ROC DE MARSAL (FRANCE)

Harold L. Dibble, Curtis W. Marean, and Shannon P. McPherron

Harold L. Dibble is a Professor of Anthropology at the University of Pennsylvania. Curtis W. Marean is a Professor in the Institute of Human Origins and the School of Human Evolution and Social Change at Arizona State University. Shannon P. McPherron is a research scientist in the Department of Human Evolution at the Max Planck Institute for Evolutionary Anthropology (Leipzig, Germany).

The use of total stations in archaeological excavation makes it very easy to record quickly the three-dimensional coordinates of artifacts, samples, or features. The greater challenge is to link these coordinates to the actual object and to maintain this link throughout subsequent processing and analysis. Such linkages are usually made through the use of unique identification numbers that are assigned to objects as they are recovered. When dealing with several thousand or even several hundreds of thousands of objects, identification numbers tend to be both long and complicated. Each can take 1–2 seconds to enter on the keyboard, and there is always the potential to make mistakes. The use of barcodes solves both of these problems: a barcode representing an identification number can be read in a fraction of a second and with a high degree of accuracy. Here we describe barcode technology and provide two examples on how the use of barcodes has greatly increased both the speed and accuracy of maintaining those crucial links and, in the process, greatly facilitated the day-to-day handling of objects.

Background to Barcodes

A barcode is a series of vertical lines (or bars) and spaces, each of which can be of different widths. Bars and spaces together are called "elements," which in turn are grouped together in different combinations to represent different characters. At a glance, all barcodes look pretty much alike. In fact, there are a number of different barcode formats that affect the way in which information can be coded within them. One of the more common barcode formats, and the one that is probably the most useful for archaeologists, is called Code 39 (or Code 3 of 9), which has nine bars and spaces: three are wide and the other six are narrow. The advantage of Code 39 is that it can represent a full range of capital letters, numbers, and special characters. There is also an enhanced Code 39 that can represent both upper- and lower-case letters. Other formats may be limited to only numbers or may not include special characters. Thus, what format you use impacts what kinds of information you can place in a barcode, though virtually all commercially available barcode scanners recognize the Code 39 format.

To some extent, the size of the printed barcode varies proportionately to the number of characters being represented. However, the size of the barcode is also affected by what is called its density, which refers to the width ranges of the bars and spaces. Lower-density barcodes have wider elements and take up more space when they are printed. The thinner the bar and spaces, the less space is required and the higher the barcode density. The trade-off is readability. Lower-density barcodes are more reliably printed and more consistently read than higher-density barcodes because minor variations (due to printing or damage) are much more serious with the latter. In addition, barcode formats have rules that specify their height relative to their width and how much space is needed before and after the barcode. In our experience, however, these constraints are rarely significant to archaeologists, as numbers in the millions can easily be expressed in a highly readable barcode of just several centimeters.

There are a number of ways to produce barcodes, although most methods fall into one of three categories. First, you can install a barcode font (free versions are available on the Internet), which allows you to use your existing software to create labels. For example, you can use a database program to print barcodes by simply applying this font when printing. The advantage to this approach is that it is free and relatively easy to customize. The main disadvantage is that you need to be very aware of the barcodes rules as discussed above, particularly those concerning sizes. In addition, there may be other formatting rules. Code 3 of 9, for instance, requires that you place an asterisk ("*") before and after the text that the barcode is to represent. Knowing this, one can, for example, include barcode printouts on a report in Microsoft Access by selecting a barcode font for the ID field, making sure that the font size meets the rules for minimum size (24 point, for instance). Adding the asterisk can be a little tricky, but one way is to write a separate query that creates a new field that might look something like this: "*" & [ID] & "*" where ID is the name of the field. When the barcode is read, these leading and trailing asterisks are automatically removed.

Another way to print labels is with commercial software. The main advantage to using a dedicated program is that it will come with a set of barcode formats and will take care of all the associated rules regarding size, format, start and stop characters, etc. These programs also typically have pre-set printing formats for printing labels available in office supply stores. The main disadvantage are cost and perhaps, depending on the software, some flexibility. One thing to look for in a barcode program is the ability to link to the database in which your data are stored. Finally, it is also possible to purchase software libraries that facilitate barcode programming. The main advantage is that you have all the flexibility that custom programming brings plus the rule-minding of the second option. The main disadvantages are that software libraries do cost money and it can take considerable time to write and debug these programs.

Many barcode formats, including Code 39, allow normal text to be written together with the actual barcode, which means that printed labels can still be read by humans. However, the real advantage of barcodes is not in their human readability but rather that they can be read extremely quickly and accurately by various instruments. The least expensive of these is called a "wand." This device looks like a thick pen or laser pointer and is passed directly over the barcode. Wands are inexpensive, but they can be a little tricky to use because the wand must be angled correctly above the label and the motion used when reading a label must be smooth and at the right speed. Thus, some practice is required to get good results. So-called "barcode guns" are easier to use but a little more expensive and a little less compact. These hand-held or mounted devices are simply pointed at the barcode, and they are able to automatically find and read the information and pass it to a computer. Barcode guns come in two types: CCD and laser. CCD guns have a digital camera inside that takes a photo of the barcode and decodes it. They typically have to be within about 6 in of the barcode. Laser guns, which are what we use, are like those seen in supermarkets (Figure 1). A beam of laser light scans the object, locates the barcode, and then decodes it. This instrument is very fast and does not need to be within 6 in of the object. In our experience, the latter are quite rugged and extremely easy to use.

Figure 1: Use of a laser barcode reader.

All of these devices typically pass the data to the computer via either a wired (serial, PS/2, or USB) or wireless (typically with Bluetooth technology) interface. Both serial and PS/2 interfaces are a little less flexible since fewer computers are built with such ports and even if they are present, with serial interfaces you may need special software to read the data. One advantage of the wired interfaces is that the barcode reader typically can draw its own power from the computer through the cable. While a Bluetooth interface requires a separate power source (such as internal batteries), its main advantage is its mobility and the fact that there is one less cable to worry about in the field. In our experience, like other wireless systems, when it works properly, Bluetooth is impressive, but getting it to work the first time can be a challenge.

One of the main drawbacks of barcodes is that they are not readable without a computer to decode their information, and over time the barcode can become damaged and unreadable through excessive folding, or by exposure to too much sun and water. It is very important, therefore, not to rely totally on them. When printing a barcode of an artifact identification number, for instance, a plain text version of the identification number should be included with the barcode. Sealing the barcode separately in a small ziplock bag is also helpful. In this way, the label will still be readable despite damage or changes in computer and barcode technology. Another serious drawback to barcodes is their potentially short use-life. The main issue in our own system is the low quality of the glue on the labels. More-expensive labels stick to bags better and longer but for exactly how long is still unclear. An even larger issue, of course, is how long the instruments needed to read them will be available. As ubiquitous as barcodes seem today, it is not unimaginable that they will go the way of 8-track tapes and record albums, and there are already a number of technologies that are perhaps superior (for example, Radio Frequency Identification, or RFID, chips), and any one of them may become the new standard. For these reasons, we consider the use of barcodes on our project as a provisional, non-archival tool designed to help us manage the collection from the point of excavation until it is turned over to a museum.

One other important issue with serious logistical implications is that the barcode system requires that each artifact is stored in its own bag, since it is impractical to affix the barcode label directly to the artifact. For us, this means purchasing 15–20,000 bags of various sizes per season. But the more difficult logistical issue comes when it is necessary to separate the objects from their bags, for example, if we want to lay out all the objects on a table. This is one reason why we immediately ink all artifacts, since there is no risk of losing their context even if they are out of their bags. However, when we are finished with such a study, potentially thousands of artifacts must be reunited with their proper bags.

Archaeological Example: Roc de Marsal (France)

Two of the authors (Dibble and McPherron, along with A. Turq and D. Sandgathe) are currently excavating the Paleolithic cave site of Roc de Marsal. Although the complete methodology is described elsewhere (McPherron and Dibble 2002), it can be described briefly as being built around the use of multiple total stations—Topcon and Leica—connected to handheld or laptop computers that collect the three-dimensional coordinates and assign unique IDs to each artifact, bucket, sample, etc. recovered from the site. In 2000, we began using barcodes to track objects from the moment of their recovery through all aspects of lab processing.

At Roc de Marsal, all artifacts over 25 mm are piece-provenienced with a total station and given a unique identification number that consists of the name of the excavation unit in which they were found (e.g., “H12”) and a sequential number unique to that unit. The combined Unit-ID (e.g., “H12-45”) represents a unique identifier for each object. The main challenge in such a system is, first, to keep the physical object in constant association with its unique Unit-ID and, second, to accurately associate any analysis of that object to this same Unit-ID. To accomplish this, we rely on two sets of labels. First are the Field Labels, which are preprinted sheets, with 24 labels per sheet, that contain new, unassigned Unit-ID numbers for objects recovered from each excavation unit. When the object is removed from the ground, it is placed in its own bag, and, after confirming the Unit-ID of the object, the next barcode label of the page is removed and affixed to the bag. These labels contain only the Unit-ID in text and barcode form (Figure 2a). In the lab,

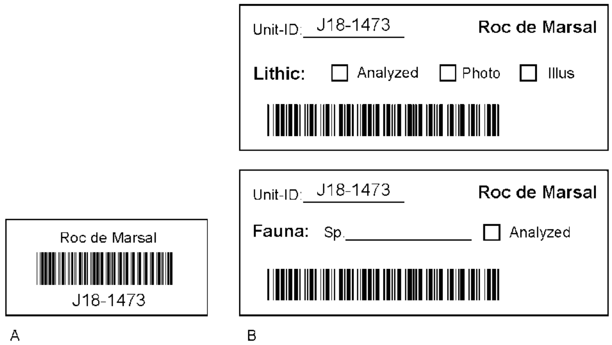


Figure 2: (a) An example of a Field barcode with only the Unit-ID; (b) Lab barcodes customized by object type.

the Unit-ID is inked onto the artifact to permanently associate the Unit-ID with that artifact, and it is transferred to a new, clean bag. The Field Label is then scanned, and a new barcode label (called a Lab Label) is printed (Figure 2b). The Lab Labels are produced for two reasons. First, by the time the object has come from the field and been through the washing-and-labeling process, the original barcode is typically in less-than-ideal condition. Second, Lab Labels can be customized to include supplementary information derived from the database and can be designed with tracking features for subsequent processing. We have separate barcode labels for fauna and lithic artifacts, and these labels include check-boxes that are used by different analysts and lab personnel to note when an object has been through the various processing stages.

The advantages of barcodes are clear in all phases of object processing and analysis. As additional data are entered for an object, such as when it is photographed or analyzed, the first step in the process is to scan the barcode. The database is then checked to ensure that the Unit-ID is indeed a valid one and that the current data have not already been associated with the particular object. If it passes these two checks, data entry is allowed to proceed. In addition, the barcodes are used any time we want to access what is known of an object. Typically this involves scanning the barcode into the GIS so that it can provide the spatial and analytical information associated with the object. Barcodes also enable us to physically re-sort material at a much faster rate—artifacts are scanned one at a time, and the computer instantly responds with the information required (e.g., level, artifact type)—an operation that takes less than a second. And finally, making lists of materials, such as objects sent out to specialists, is equally fast—the labels of the objects are simply scanned into whatever software is used to keep the list. The use of the barcodes to enter the Unit-ID numbers is many times faster than doing it by hand, and much more accurate.

Archaeological Example: Mossel Bay (South Africa)

The Mossel Bay Archaeology Project (MAP) is a long-term field study of the Middle Stone Age in the Mossel Bay region of South Africa (Marean et al. 2004). MAP has been using total stations as its primary means for measuring and recording field observations since the first field season in 2000, and starting in 2001, we integrated barcode scanners into our total station-based recording systems. In this project, all finds, including lithics, fauna, and any other artifact or ecofact, are piece-provenienced (which we call "Plotted Finds") with the total station, which is also used when drawing features and stratigraphic profiles. We also shoot "chits," which are small paper targets, on all plan and stratigraphic photographs so that they can be rectified and built into our GIS. Days can go by without anyone on site using a tape measure, and no tape measures are used for any drawing or measurement recording, due to their inherent lack of precision (McPherron et al. 2005).

Between excavation and field mapping, massive amounts of total station measurements are taken—tens of thousands per field season. Tying these measurements to their context could be a daunting task, both in terms of the mechanics of recording (typing field descriptions into a total station or hand-held computer) and the potential for error in that typing. Our combination of barcode scanners, surveying software, and database management has overcome this problem.

The typical MAP excavation uses two total stations positioned at opposite ends of a cave. Each total station is positioned to cover a specific area being excavated during that day (Figure 3). Our total stations are reflectorless and are operated by hand-held computers running software called Survey Pro (developed by Tripod Data Systems, or TDS). Most total station brands have reflectorless models that cost only a thousand dollars more than those that require prisms, and we have found that the capability is cost-effective. Although we often use chits (Figure 4), with adequate light one can shoot directly onto anything. Thus, reflectorless systems allow one to shoot onto and map features that are dangerous or difficult to access. Based on our experience, the reflectorless units are as accurate as those requiring a prism within ranges from 3–350 m.

In the field, we use two types of barcode scanners because of differences in the two field computers that are used by the project. The first is a wired scanner (HHP Imageteam 3800) that is powered by and communicates with the handheld computer (a TDS Ranger), which, in turn, is cabled to the total station. These field computers are rugged, and we have not had a single failure over four excavation seasons. The Ranger comes with a keyboard wedge program that will receive the barcode data and translate it as regular keyboard entry, and we designed a hardware interface that allows us to attach and power the barcode scanner to the Ranger-powered serial port (Figure 5). While wiring your own hardware interface may seem intimidating, it is actually quite easy, as the pin-out codes are provided in the Ranger and barcode-scanner manuals. If necessary, professional electronics technicians, either at one's university or a commercial dealer, can build the interface for you cheaply. Our other field computer, however, a TDS Recon, does not have a powered serial port. The solution was to add a Bluetooth card to the Recon and then purchase a Bluetooth-enabled barcode scanner. Several varieties run on batteries, and we use one made by Socket. These appear somewhat more fragile than the HHP, but we have gone three full seasons without malfunctions. Ultimately, wireless scanning may be the most flexible solution, as it will allow the barcode reader to move independently of the handheld computer.

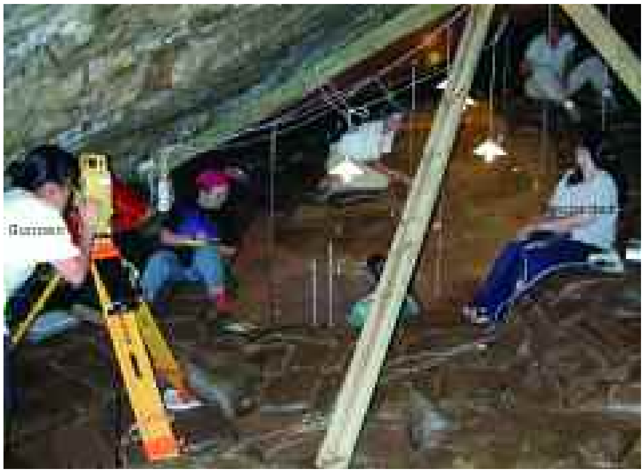


Figure 3: The set-up in an excavation area, showing the total station being run by the Gunner (left), the Recorder with the hand-held computer and barcode scanner (right), and the excavators.

Figure 4: A plan photograph of a feature (a burned area) showing targeting chits that will be shot by the total station. This particular set-up is for a photograph that will then be rectified and built into our GIS.

Figure 5: The Recorder with the Ranger handheld computer and the barcode scanner, scanning a Plotted Find and Lot Number to Survey Pro on the Ranger, thus linking the find to its x-y-z coordinates and stratigraphic provenience.

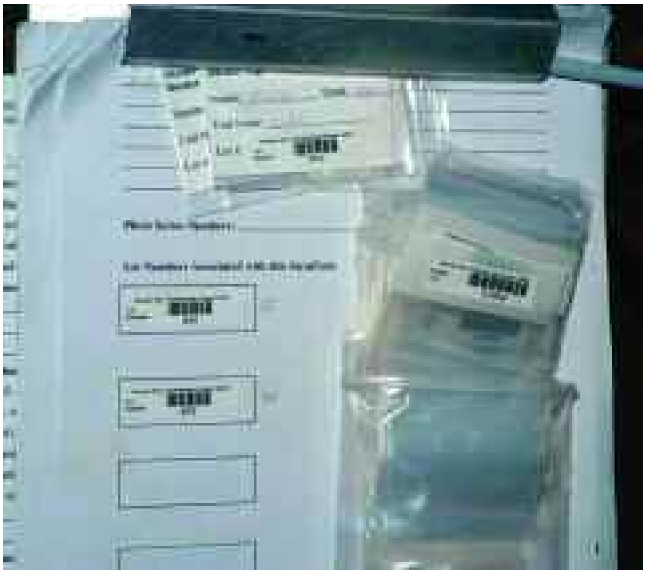


Figure 6: A set of Plotted Find bags showing the barcode labels, and Lot Number labels affixed to the form describing that StratUnit.

The challenge of any complex project is to integrate all of the measurements and descriptions easily and without long and repetitive data entry. To accomplish this, we use a relational Access database anchored by three key tables that share two fields—the Plotted Find number and a Lot Number—that define the relations among the tables. Plotted Find numbers are sequential, and so every find or sample receives a unique one. Using software called "Bartender,™" the barcodes and the readable text are preprinted on sheets of return address labels. These labels are then put inside small ziplock bags and distributed to excavators in batches in small boxes (Figure 6).

Like many archaeological projects, our excavations take place within a horizontal grid system (1-m squares, and within them, 50-sq cm quads) and some type of stratigraphic unit (what we call a StratUnit) that reflects an observable anthropogenic, biogenic, or geogenic observation. Thus, all excavated materials, including Plotted Finds, will be provenienced minimally by StratUnit-Square-Quad. To each set from this provenience, we assign a single identification number called the "Lot Number," which serves like a tracking number for a package. We generate in advance a second set of barcode labels for these Lot Numbers for use on bucket tags and excavation forms (Figure 6). When an excavator begins to excavate a new StratUnit-Square-Quad, they add this to the Lot Number table. This means that any find or observation that has a Lot Number attached to it can ultimately be associated to its full StratUnit-Square-Quad provenience through a relation to the Lot Number table. This also means that this is the last time that anyone needs to enter or write the StratUnit-Square-Quad—one can now scan the Lot Number barcode to get that information.

The process operates as follows: an excavator puts one or two chits in place of the artifact, and drops the artifact into a bag with a barcode Plotted Find label. The "Gunner" (the person running the total station) targets the chits, and the "Recorder" (the person who runs the handhelds and keeps notes) activates the total station from the handheld. The shot is taken, the excavator passes the bag to the Recorder, and the Recorder scans the Plotted Find number and then the Lot Number to the Survey Pro database; this ties together the x-y-z coordinates, the find, and its provenience. Similarly, when mapping a hearth, the excavator places several chits, and for each shot, the Recorder scans the Lot Number, tying the feature being shot to the multiple x-y-z coordinates. Survey Pro can be set to carry over previous entries to the next, and thus if desired, multiple shots to items or lots with the same lot number need only scan it once.

Conclusions

This brief note described the use of barcodes in on our own excavation projects and also presented some background and issues surrounding their use. It is clear to us that barcodes are very effective in our own projects and that the expense of this technology is not prohibitive since it is offset by the vast increase in speed and accuracy of entering information. Although circumstances vary widely among archaeological projects, the use of barcodes is something that all projects should consider. One final point: while we describe here the specific hardware and software that we are currently using, computer and instrumentation technology is constantly changing. Before making any purchases, users should consult with dealers to make sure that all of the various components will work together smoothly.

Acknowledgments

The Roc de Marsal project is supported by the National Science Foundation, the Leakey Foundation, the University of Pennsylvania Research Foundation, the University of Pennsylvania Museum of Anthropology, the Service Régional de l'Archéologie, and the Conseil Général de la Dordogne. The fieldwork at Mossel Bay is supported by the National Science Foundation, the Hyde Family Foundation, the Institute for Human Origins, the Dias Museum of Mossel Bay, and Iziko South African Museums.

References Cited

Marean, C. W., P. J. Nilssen, K. Brown, A. Jerardino, and D. Stynder
2004 Paleoanthropological Investigations of Middle Stone Age Sites at Pinnacle Point, Mossel Bay (South Africa): Archaeology and Hominid Remains from the 2000 Field Season. *PaleoAnthropology* 2:14–83.

McPherron, S. P., and H. Dibble
2002 *Using Computers in Archaeology: A Practical Guide.* McGraw Hill, New York.

McPherron, S. P., H. Dibble, and P. Goldberg
2005 Z. *Geoarchaeology* 20:243–262.