The optician's trick: an approach to recording excavation using the IFPRS

(Iconic formation process recognition system)

Max Adams

'How can it be measured? It cannot be measured. It is a notion; a most valuable notion, I am sure; but, my dear sir, where is your measurement? It cannot be measured. Science is measurement - no knowledge without measurement'.

(Patrick O'Brian, *Master and Commander* 1970)

Introduction

Processualist excavators and post-processualist theorists have been doing their own thing in isolation for so long now that it is hard to recall that as long ago as 1990 there were voices of reason calling for method in spite of rhetoric (Bell 1990). Now the theoretical mainstream has finally thrown down a gauntlet, challenging the discipline to put its methodological money where its mouth is (Hodder 1997). A reflexive archaeology is called for: we must be critical of our assumptions, we must be relational, we must be interactive, and we must be responsive to multivocality (*op cit,* p.694). These are laudable aims, and recent literature is full of similar pleas to throw off the shackles of reductionism.

However, as Chadwick (1998) has made very clear, while all eyes have been focused centre stage on *The Great Debate*, voices offstage have been busy, and for quite a while, with their own discourse. Ceramicists, micromorphologists, environmental analysts and other consumers of primary archaeological data have become increasingly demanding and vocal in their desire for more sophisticated and interesting data, for interaction, and for the right to set research agendas (for example Cumberpatch and Blinkhorn 1998; Barham unpub;). Excavators have begun to reject the notion that they are or should be passive observers of archaeological entities and, indeed, to reject the notion that archaeological data are themselves passive.

There is an increasing awareness affecting all levels of field enquiry in British archaeology that, above all, immediate on-site recognition and interpretation of evidence for formation processes is the key to improving the quality of inferences made from excavations (a brief glance at the Interpreting Stratigraphy series should demonstrate this). The formation processes movement, inspired by the work of Schiffer (1987) and other behaviourists was launched in Britain more than a decade ago at the Theoretical Archaeology Group conference (Bradford 1987). It was followed up by the author (Adams 1991; 1992a; 1992b) and others (e.g. Boddington, Garland and Janaway, Eds. 1987), and there is now a purely British formation processes bibliography with several hundred entries (Adams, in prep). Archaeologists themselves are increasingly seen as formation processes (Kristiansen 1985, Adams 1989). It is astonishing, then, that Hodder's state-of-the-art method requires a social anthropologist to point out to excavators at Çatalhöyük that *'we did not think through how artefacts had become incorporated into the archaeological record'* (Hodder 1997, 698). Given this statement, it is hard to see how archaeologists themselves are to be incorporated into a

In Roskams, S. (Ed) Interpreting Stratigraphy: Site evaluation, recording procedures and stratigraphic analysis.
BAR International Series 910: 91-102

formation process methodology except in the sense, noted by psychologist Steven Brown (1997) of subjectivity as error variance.

What follows is an account of how, over the past seven years, a system has been built in which explicit recognition of formation process traits provides the core for an excavation method which demands that both data and archaeologists are seen as dynamic, complex and non-linear. It has not been built in isolation. Workers in many disciplines from political science and psychology (see Brown 1997; Webster 1997) to criminology and public policy, are aware of the problems of dealing in subjective, 'fuzzy' data derived from human witnesses, and they too have been developing techniques to build subjectivity into method rather than, like Binford two decades ago (1977, 404), attempting to deny it. Archaeologists are not the only social scientists to suffer from physcis envy.

Context (single)

The adoption of the broadly standardised single context recording sheet by virtually all professional archaeologists in Britain has ensured that data from widely differing archaeological environments can be compared. And yet, some field archaeologists have an uncomfortable feeling that such systems fail to deal with types of information which are both complex and interesting: subtleties in the depositional record which would, if only we could get at them, reveal deeper truths. In our drive to be objective and scientific we have fallen into the reductionist trap of defining our interests by our ability to deal with them. We have opted for a spurious conformity rather than a plurality of approach, so that we can all say that we produce the same information. In other words, standardisation has led to mediocrity - error free mediocrity, perhaps, but mediocrity all the same.

For most of the 1970s field archaeologists were keen to please the proponents of New Archaeology: field recording, because it was recognised as being very difficult, had to be extra rigorous in order to be scientific (Binford *op cit*). Objectivity was the key. Hence the recording systems of that era, some of which are still with us in spirit (for example the current English Heritage CAS` manuals), are extremely detailed in their requirements for measurement and objective observation. Later, in the 1980s and 1990s, a seemingly insuperable void appeared between theorists telling us that archaeological truth is relative (e.g. Shanks and Tilley 1987), and data managers telling us that we still had to behave 'as if' we were being objective and scientific (HBMC 1991). How could archaeologists do science while viewing themselves as a sociological phenomenon? At the same time the fractal (cf. Gleick 1987) nature of archaeological resources was becoming apparent: the closer you looked at the soil, the more complex it became (and the same might be said of the archaeologist).

An obvious implication of this unease was that we were going to have to adopt recording systems that were able to cope with more detailed and complex observation in the field. In a mostly competitive contract market this posed a problem, because it sounded as if field recording, in order to adequately reflect theoretical demands, was going to become more expensive. At the same time, we had to deal with the paradox that a discipline which was becoming hungrier for complex data sets was having to deal with the realisation that objectivity was, perhaps, an unobtainable (perhaps even undesirable) goal (Adams and Brooke 1995). How, then, to be increasingly scientific when we had to rely on inherently 'flawed' subjective observation? What could field archaeologists do with their systems to deal with this paradox, without it literally costing the earth?

Not many archaeologists have the luxury of setting up a unit from scratch. In 1993 when *Archaeological Services* were set up at the University of Durham there was a rare opportunity for the author to experiment with radical solutions to the objective/subjective dilemma. It was felt that a commercial unit working within an academic framework ought to approach the problem from a research-oriented perspective, but it was clearly crucial to find ways of reflecting archaeological complexities without pricing the unit out of the market.

The optician's trick

An oddly analogous discipline seemed to provide some conceptual as well as technical answers. Opticians manage to provide their patients with accurate and reproducible prescriptions for their eyes which, in the vast majority of cases, result in artificially improved vision. Despite the complexity of our visual equipment, opticians are able to accurately assess the nature and extent of visual impairment by asking us a number of questions which are purely subjective *(is this sharper than that; is the red circle clearer than the green?)*. The patient requires no training in this technique. Better still, we are using flawed eyesight to correct flawed eyesight. The simplicity and effectiveness of the technique reflects the fact that human beings are much better at perceiving change than they are at perceiving quantitative phenomena. For example, we are quite adept at detecting temperature change but rather poor at guessing temperatures in degrees Celsius or Fahrenheit; we notice when part of our environment has changed, even if we cannot quantify precisely what has changed. All field archaeologists will be familiar with the Munsell problem, and the criminal justice system has spent a great deal of energy trying to ensure that witnesses are relieved of relevant information without being led: there is a whole world of difference between asking: *was the accused wearing darker or lighter coloured clothing?,* and: *might he have been wearing blue tracksuit trousers?* The equivalent 'objective' question would be: *please describe the scene of the crime,* first making sure that the witness had been trained to recognise salient features of the crime scenario. In archaeology, as in criminal investigation, the crux of the problem is to identify processes of formation by interrogation of a witness, either directly at the time, or indirectly through a data set. Archaeology has the advantage over the criminal investigator that it can train witnesses beforehand.

Would it, therefore, be possible to design a system which utilised the field archaeologist's powers of subjective perception to produce a record of equal or greater integrity than the standard objectivised system? Clearly, as the optical and criminal analogy shows, any query-oriented system must involve very carefully prepared questions. Would it produce accurate information and would it, perhaps more importantly, produce more interesting information, reflecting our desire to capture more of the complexity of archaeological systems? Such a system would have to be highly data efficient, and at the same time be capable of interrogation at quite subtle levels, as well as retaining compatibel metric observations where appropriate. For example, in describing the characteristics of a post hole void, measuring its depth and diameter and drawing its plan and profile is till, for all the problems of using measuring devices, better than asking the excavator whether it was more like one post hole than another. Digital 3-D mapping would be better still. In addition, since it was recognised that archaeologists themselves are a vital (and not to be underestimated) part of the archaeological process, there would have to be a way to account for their behaviour. For example, what is the response of an excavator to heavy rain, or an intractable interface on a Friday afternoon? Is the tendency to overrationalise, or to fail to investigate or record fully?

Formation processes

In order to have a chance of success any system had to involve the explicit recogniton of formation process traits on site whilst allowing fruitful interrogation later. This is because all archaeological inferences are based ultimately on either assumptions about or observations of such processes, since the primary unit of archaeological information is the trait formed by an impact of human on artefact or ecofact (Adams 1991). The dangers inherent in failing to recognise formation processes at work have been exposed already (Schiffer 1987; Adams 1987), and increasingly specialists in material culture and environmental data are demanding much more explicit information relating to the processes by which material comes into their hands. The training of archaeologists in environmental and other specialist on site techniques is to be wholeheartedly praised, but it has profound financial implications. It was necessary to build a system with the inherent ability to train its user as it progressed, so that, in a computing analogy, the excavator became part of an expert system into which he/she fed back new recognition skills as they were developed.

As is often the case, a solution was suggested by a problem which had arisen during post-excavation. The author had been asked to prepare a cemetery excavation for archiving and publication (Adams 1996). The cemetery in question, at Addingham in West Yorkshire, had been recorded using a standard single context system which offered no specific tools for recording inhumations. Most of the excavators, though highly competent, were inexperienced as far as cemeteries were concerned. It became obvious that some depositional aspects of the cemetery were quite complex. Bones had apparently been transferred from one grave to another, and there were widely differing preservation characteristics within the bone assemblage. There were also grave shaped pits with no visible bones in their fills. Were these to be classed as empty graves and, if so, what were the formation processes which led to the absence of skeletal material? Much of this information had been recorded by the excavators in the only way open to them: as free text in a description field. Their observations had been acute, if inconsistent, but the information was unusable in its original format and there was clearly room for a higher level of intepretation than that which was obvious on site.

It was decided to effectively re-record these observations using a presence/absence field in a dBase IV database. The questions posed were explicitly formation process oriented, aiming to identify regularities and irregularities in the depositional characteristics of each grave. Terms such as 'well-' or 'poorly preserved' were of little use, but qualitative language, such as the common description of bone as 'crumbly, with little white crystals' allowed the interpretation that bone in some cases had been reduced to brushite (mineralised), with its attendant implications for reconstructing the depositional history of the site. The problem of how to statistically analyse combinations of characteristics was only later, and crucially, addressed in the Durham system. Nevertheless, the exercise revealed an underlying behavioural pattern on the site, with graves which lay closer to the church being favoured for burial and reburial. There was a positive correlation between bones which had been severely mineralised (despite the well-drained nature of the soil) and those which were being deposited after disarticulation had occurred -in other words, the bones had been disinterred and moved to a grave closer to the church, causing the mineralisation process to accelerate.

It had been fairly easy to determine a series of questions to pose of the Addingham archive, since the information which needed to be analysed already existed. Was it possible, though, to identify a series of questions to ask of field archaeologists which would deal with a wide range of sites and depositional scenarios; which would, furthermore, be meaningful during the post-excavation process. All the archaeologists who worked for the Durham unit during its first two years, as well as a number of specialists, were instrumental in identifying and then modifying the questions; these were experimentally introduced for an urban site which, happily, produced more than its fair share of inhumations (Carne and Adams 1995).

It was decided, firstly, to arrange questions into discrete groups relating to different aspects of archaeological information. These questions are posed in the form of icons (Figure 1), pictures of simplified scenarios which remind the excavator of a textual question written in the manual. A series of nine questions (nine is a good number, for reasons which will become clear; there is a historical analogy here with Gregor Mendel, who chose to select for seven characteristics of peas, which fortunately have seven pairs of chromosomes: (Bronowski 1968)) asks the excavator what the horizon that they are dealing with is like: *is it easy to recognise, impossible to recognise, is it weathered,* and so on. In combination with other sets of observations, much can be inferred concerning the formation of a deposit and subsequent impacts upon it from the nature of its boundaries - this goes for vertical interfaces as well as deposits. If the site is characterised by ill-defined interfaces what does this tell us about its formation, and what are the implications for the inferences that will be built with such data (see figures 2 & 3)?

A second group of nine questions asks the excavator to specifically identify, and justify their identification of, traits of processes which can generally be easily identified during excavation: *is there evidence of animal or plant disturbance; is there evidence that a deposit was formed over a short or long period; is there evidence that the deposit was formed by wind or water,* and so on. The answers to these questions are combined with those concerning the nature of the horizon to develop a model of the formation characteristics of individual contexts, and later groups of contexts and the whole site. For example, a flat rural site is often characterised by negative features cut into subsoil. *Are all the features truncated, or have some been protected by deposits of loess; are there regularities between some of the feature fills; does the finds distribution reflect the location of processing activities, or are the finds predominantly located as a result of secondary discard and later reworking, for example by ploughing?* These can be supplemented by later analysis of micromorphology, for example.

Excavators who are prompted to ask detailed questions about formation processes become extremely good at recognising the evidence for them, enhancing their own levels of skill in the process, producing data which can be analysed, and yet not having to suffer the agonies of filling in endless free text fields and not knowing how to say something interesting, even if there are interesting things to observe.

There may be more initial problems with the third set of six questions, which relate to material culture retrieved from deposits (for the actual prompts, see below). Why are these questions being posed? There are two reasons. First, it is important for the excavator to be thinking about the circumstances in which a deposit was created. *Can you tell if it was it as a result of a manufacturing or processing activity? Was it as a result of maintenance or discard processes? Does the artefact carry traces of use wear or reuse?* Sometimes, though not always, artefacts may provide evidence of such processes. Better that these are observed at the time of retrieval, so that the information can be fed back into the excavation process. Secondly, such information about material culture is usually taken out of the analytical system (when finds go to their specialists) and is only reintegrated at a late stage of reporting, a fact which has been critisised by many workers over the years. In other words, that information often does not inform the process of analysis itself. In the case of material culture the icons are synthetic, in that they do not request identification alone, but also interpretation, and that interpretation is overtly based on the work of Schiffer (1987). This view may be unattractive to those who see a clear difference between information gathering on site and interpretation off it, but it has the advantage of being explicitly defined, and calls for demonstration of evidence when an artefact is interpreted as other than secondary refuse.

Some excavators are uncomfortable about making judgements of this nature. So the system actually asks the question: *can you, the excavator, determine these characteristics with confidence?* If he or she can't, that is effectively recorded as the tolerance inherent in that archaeological context. The reasons for the indeterminate nature of the problem are the sum of the processes which have operated. The evident problem, that excavators learn as they do and that all will not be equally adept at identifying such traits, leading to a data set of questionable use, will have occurred to readers as it occurred to the authors. The answer to that problem is inherent in the system, and will be discussed below.

The fourth set of questions is as soul-searching as even the most die-hard post-processualist could desire. Six further icons prompt the excavator to assess the process of excavation and recording which they have conducted, and highlight problems which they believe may compromise the inference potential of the archaeology. The idea which is propagated among unit members is that being wrong or having doubts is okay (in moderation), so long as we know why and in what respects. There are fields for lousy weather, time constraints, general uncertainties, and an 'I screwed up' field. Since these questions relate to crucial aspects of the process of doing archaeology, they have an equal status with the more traditional observations, and have to be qualified with details.

It is true that some recording systems have attempted to address aspects of this problem with fields such as "risk of intrusion" (Steve Roskams pers. comm.). The very word "risk" exposes much of the underlying philosophy of such systems: it suggests that intrusion (whatever formation process that is supposed to reflect) is a contaminant of an otherwise pristine record. What the Durham system would seek to identify, rather, is the probability of material - artefacts and ecofacts - identified in one depositional event having migrated from another after deposition. The probability of this event would be judged by examining the integrity of the interface between the two (or more) deposits. For example, root fibre penetration between deposits is common, especially on flat rural sites. The probability of particles of the diameter of a root fibre migrating between deposits must be seen as high, at any rate high enough to be flagged as a problem if the deposits in question prove to be crucial for the interpretation of the site. A three inch diameter sherd of pottery is mnost unlikely to migrate along the path of a root fibre. However, a series of deposits whose interfaces have been compromised visibly by animal burrows would prompt a query about the integrity of the same sherd. Interfaces between deposits where the excavator has trouble defining precisely where the boundary lies, a common scenario, may suggest substantial floral and faunal penetration in the past without current visual confirmation; the integrity of the deposits must still be suspect. It is woefullyinadequate to pose the question "what is the risk of intrusion?", without characterising the depositional and interfacial scenario which is supposed to present the so-called risk. A simple illustration is shown in figures 2 & 3, in which the same section has been drawn in two ways.

The first, conventionally, shows deposits separated by interfaces which are defined by the width of the pen nib used, to rationalise the boundary. What information is the informed skeptic being given to make his or her own judgement of the integrity of the deposits -let alone the integrity of the excavation process? The second shows the same section with interfaces defined by iconic reflections of processual icons: an instant visual guide to the interpretive potential of the record. The moral is that archaeologists need to stop employing a mindset that sees some processes as good and others as bad (as in 'badly preserved'). This idea should have been laid to rest a decade or more ago (Boddington et al 1987, preface).

Data collection

On the face of it, the field archaeologist is being asked to do a lot more work in order to reflect the increased need for complex observations. This problem is obviated by the data capture technique used in the system. Although excavators are expected to remember the precise nature of the questions being posed, they are prompted on the context sheet by pictures, or icons, which remind them of the question. All they have to do is circle a number under the picture to indicate yes, or leave it alone to indicate no. In that way, answers to the thirty extra questions they are asked can be recorded very quickly (quicker by far than filling in free text). Qualifications to some answers are required, and these are noted in a deliberately small free text field. If this field appears to be over-used there may be a case for adding or changing icons, so there is constant implicit feedback. Excavators get faster at answering these questions as their perceptive skills are enhanced and they become more able to recognise traits more quickly. Readers, again, will object that there is a danger that repetition of use will lead to sloppy recording by excavators, or that they will increasingly record similar sets of observations for contexts which have overall similarities. Excavators are, however, reminded that analysis of their sheets will detect anomalous decreases in the diversity of information - in other words, one can tell if the system is failing.

How is this done? Reference to Figure 1 will show that beneath each icon is a number. Many of the icons are self-explanatory, but the relevant sections of the crib sheet are included below for interest (it is possibly the briefest recording manual in use, although it does rely on an understanding of the more objectivised fields). Let us say that, in the horizon/interface section, the observations 1 and 16 are made (note that some of the icons are, or ought to be, mutually exclusive: this can be a useful error-checking device, but sometimes, gratifyingly, apparent contradictions actually point to characteristics which had not been anticipated). During excavation (which means getting instant analytical feedback) or post-excavation the numbers are added together, producing 17. The set of characteristics relating to horizons, stored as the number 17, can only be those two icons which have been ticked. The same goes for the other set, so that each set of nine questions has a computer field consisting of a number between 1 and 511 (if the number is actually 511 you need to sign up the excavator for retraining, but you could, of course, have the computer flag some numbers which, when entered, automatically produce an error/confirmation message; for example, a score of 9 under the horizon field would indicate a horizon which was both perfectly clear and visible, but which was also indeterminate). For sets of six questions the same process applies, except that the maximum number of combinations for each is 32 and 64 respectively. In the last set, any number between 0 and 63 is plausible; this time, if the observed result leads to the number 63 going into the database, then you are being told that for a variety of reasons the site probably shouldn't have been dug. Sets of more than nine characteristics produce irritatingly big numbers, and it seems that more than nine options make visual scanning of the icons difficult. This ought to be put to the test in a forthcoming experiment with an eye-tracking device, which will seek to evaluate the system as a visual process.

Anyone with a calculator bigger than that belonging to the author will be able to work out exactly how many possible combinations of observations can be recorded in this way - it is something in excess of 534 million per context, in any case. This constitutes an awful lot of information. In order to understand what this system may be able to offer in terms of generating inferences it is important to separate its potential applications from the question of the actual analytical techniques which are applied to it. In other words, the author can think of many interesting questions to ask of the information, but has yet to develop the tools required to indulge in such analysis.

Potential for analysis

It may be rather trite to suggest that the possibilities are endless. Nevertheless, it is difficult to see how the analytical potential could be constrained except by imagination and software. The most important aspect of this information is that it highlights clusters of characteristics, so that it is the combination of processes which can be interrogated. There are many levels of fairly obvious analysis which can be undertaken. To begin with, a simple query like *how many ditch fills contained residual pottery and had been formed by weathering* may point to inferences concerning the nature of assemblages and their origins. On flat rural sites it is relatively easy to determine whether material culture relates to structure function, because a complete lack of primary refuse, for example, will suggest that finds distribution relates to maintenance or abandonment activities, or to later reworking by ploughing. This sort of level of information becomes available while excavation is in progress, and informs that process. Later, unobserved but not necessarily less important clusters may emerge from statistical analysis, pleasingly fulfilling Clarke's prophetic statement in 1968 that archaeological entities should display surprising, 'emergent' behaviour (1978, 70). Indeed, it is the experience of the Durham team that even inputting iconic information onto a database provides a good initial grasp of the range of characteristics.

More interesting, perhaps, is the power of the system to assess the quality of inferences which may be generated. Indistinct horizons, widespread root or animal penetration, lousy weather and questionable relationships are all, as every excavator knows, likely to constrain the potential for inference. To be able to analyse the evidence for this is a new departure. Conversely a feature, structure, site or even site type with characteristics such as clean horizons, lack of contamination and finds in important contexts, can be demonstrated to have an inherently high potential for interpretation. Evaluating both site types and archaeological resource bases by this method could prove a very powerful tool, and would provide a stronger and more explicit basis for assessment phases such as those required in current archaeological project management paradigms (e.g. MAP II: HBMC 1991), than many of the questionable, often implicit, techniques being used at present. Phrases such as "self-evidently important" drive this author to distraction.

Even better, especially from the point of view of post-processualists, is the potential for analysing archaeologists themselves. What happens when analysis shows that a site where there are strong clusters of formation process characteristics has a number of features which show radically different characteristics? Is there a set of processes at work which may lead to a fresh interpretation of that part of the site (as, indeed, happened at Addingham)? Or do the anomalous clusters coincide with the work of a single excavator? If the latter is the case, the implication is either one of both retraining and of viewing the suspect data with informed scepticism, or alternatively demonstrating how and why the excavator is performing differently from his or her colleagues.

It has already been the experience of the author that there is a common confusion between secondary and residual refuse, and this has proved very useful in prompting the author to phrase questions more carefully and explicitly. More subtle information concerning the way in which excavators' behaviour changes due to archaeological and non-archaeological influences can be looked forward to in the future. In the meantime the system has proved valuable as a teaching aid, enabling students and professionals to evaluate their own perceptive and interpretive skills. The way in which the system mediates between the archaeology and the archaeologist by way of a visual grammar is now the subject of a study by psychologist David Webster of the University of Durham (Webster *in prep*), whose interest in the system reflects strong academic concern for the ways in which humans impose order (or indeed chaos) on their visual environments, as will be seen below.

Techniques of analysis

In the mind's eye there is a picture of little bits of information clustering together to form a multi-dimensional galaxy of problems and answers which relate to both trivial and deep aspects of the formation characteristics of an archaeological site. In practice this may be rather difficult to achieve, not because of the potential size of the data set and the number of relationships involved, but because of the unspecified nature of the relationship between each of the characteristics. It is true that some of the icons should be mutually exclusive, and that there are some fairly obvious clusters which one might expect to retrieve, but there is no embracing logical imperative to the scheme (this is seen as a conceptual strength by the author, who may be alone in this). To use an analogy from modern physics, the analysis should be able to map the entropy (tolerance would be a better term) of an archaeological system. The more detail that we see within the entropy, the greater the opportunity there is to ask complex questions: the more 'interesting' the problem becomes (see Gleick 1987, 241-73). Perhaps a better analogy would be the development of a neural network, in which the strength of links between nodes with initially unspecified relationships reflects deep characterisitcs of a dynamic system. There is an echo here of the long running dispute in psychology over Q methodology (Brown 1997); interestingly, both William Stephenson (1988), the British physicist/psychologist and the author openly acknowledge their debt to Werner Heisenberg, whose tolerance principle (1927) demonstrated that limits to knowledge were set by the observer.

It has already been noted that pre-determined combinations can be proscribed during database entry. Other combinations, however, appear with frequency. For example, on a rural site comprising major ditched enclosures, there is quite often a co-incidence of fills which are silty, bear evidence of having been laid down by water, have no root or vegetative penetration, and are laminated with clear, even interfaces. Artefacts and ecofacts from such deposits would generally offer a high level of interpretation for dating episodes in ditch histories. Such combinations are easy to pick from a database and allow specialists to treat material from such contexts in a different way from material retrieved from secondary or residual contexts. Indeed, any specific query of that logical variety can be handled quite easily: *do the most chronologically diagnostic finds come from contexts which are depositionally reliable?* Search for deposits and horizontal interfaces with clean horizons, little evidence of reworking, an absence of secondary deposition and no animal or floral disturbance. In the Breamish Valley in Northumberland the system has been used to identify which carbon samples are likely to offer the highest potential for dating significant landscape events (see below).

More ambitious questions require a slightly different approach. One would almost certainly want to interrogate the data to determine if individual excavators are behaving in a consistent way, that the diversity of observation changes because of perceptive improvements rather than boredom (although it may be that boredom produces the opposite effect: perhaps the psychology literature contains an answer). Such queries can be dealt with by the application of tools from a system like SPSS (Statistical Package for the Social Sciences). However, one may wish to see a more visual output reflecting highly complex relationships, so that the processual fingerprint of a site would appear as a distinct and unique pattern of clusters: a sort of fuzzy chemical molecule with bonds strengthened by the persistence of correlations. Such visualisation may, ultimately, allow deep understanding of some archaeological phenomena which, until now, have mainly been dealt with intuitively. The author believes that neural networking may be the most profitable form of analysis in the future. For the present, Jane Gleghorn of the University of Newcastle has prepared a data entry package in Microsoft Access. Although based on a standard relational database, the data entry screen features the icons from the recording sheets, as buttons, which are pressed to record the presence or absence of that observation. The numbers are automatically added to produce a numerical combination which will shortly, it is hoped, be fed into a suitable analytical software package. Data entry is, incidentally, quicker than for a traditional single context system; the author can enter 25 full sheets per hour. It would be fun as an exercise to employ the full system to reinterrogate data sets which have been recorded using other systems, as was done with the Addingham data. What new relationships would emerge?

Applications in rural excavation

During and since the development of the system *Archaeological Services* have excavated two major rural sites: a large late Iron Age enclosure and settlement at Port Seton, East Lothian (Adams and Haselgrove 1995), due for destruction by a housing development, and elements of a multi-period landscape at Ingram Farm in the Breamish Valley, Northumberland (Adams and Carne 1995-98; Adams 1999). Both projects involved the integration of a team of professional excavators with undergraduate students from the University of Durham. All students were trained in the use of the Iconic Formation Processes Recognition System (IFPRS), in many cases having never used any other form of single context recording, and often never having held a trowel before.

With few exceptions the system has succeeded as a training tool. With only three weeks experience students acquire a visual and analytical grammar and vocabulary of surprising sophistication, even though technical and motor skills often take much longer to develop. There has been some scepticism from professionals used to more traditional systems: many dislike the deliberately small free text description field until they become comfortable with the combination of icons and qualifying testimony. It is certainly appreciated for its speed of use and there seem to be few problems with interpreting the iconic queries. Excavators who have become involved in the post-excavation process appreciate the explicit and easily retrievable nature of the processual information, and quickly rely on it to grasp a sense of the overall depositional and post-depositional characteristics of a site. Students who have gone on to work as professional excavators have found themselves shocked that most single context recording systems fail to allow for such complex data gathering.

Specialists such as faunal remains analysts, soil and plant experts and ceramicists find that the information which they would ideally like to support their research is not only full but immediately available after the initial error checking phase. This means that there is a far more meaningful dialogue with specialists than was hitherto possible. For example, the evidence of quality of a sampling environment, and the conditions in which it was taken, as well as stratigraphic and spatial context, can be supplied simply by inclusion with the material of the relevant context sheets, or statistical summaries of the iconic information. Questions asked by specialists which have not been met by the system have often been incorporated into the next updated version. Each archive includes a copy of the version used to record that site. It may be argued that this compromises inter-site comparability. The author would argue that inter-site comparisons made on the basis of data compatibility are bogus, since the compatible nature of the data set surely masks incompatible techniques of data recovery and interpretation.

In the Port Seton archive *(op cit)* an account was given of the overall characteristics of the site in terms of cultural, environmental and operational processes, information for which was drawn from a simple statistical scan of the iconic clusters. This account made it clear that the information potential of the site was constrained by a general lack of primary or *de facto* (ie. abandonment) refuse, by the small size of the sample of ditch fills, and by truncation caused largely by later ploughing and, in one case, by over-machining (the over-machining was recorded on context sheets as an operational impact, so that there was no question of supressing information from a site diary or a conveniently edited video recording). The strength of this approach is that it provides explicit evidence of ways in which the archive can be interpreted, and in some cases a warning of how it should not be interpreted. On another site it emerged from analysis that one of the most influential destruction processes had been over-zealous washing of pottery. This criticism was included in the archive and in published summaries.

At Ingram Farm in the Breamish Valley the strength of the system has been demonstrated by the quality of the dialogue which it has engendered, and by its role in developing new excavation techniques for use in very difficult soil conditions. Students in their first year of study are quickly empowered to develop their own observational skills and as a consequence they rapidly gain the self-confidence to challenge assumptions (e.g. that the soils of the Cheviot Hills are all leached podzols) which they have been warned of in the classroom. Second year students who return to Ingram have noticeably improved skills, particularly in the way in which they are able to articulate problems of perceiving and understanding archaeological phenomena. It has also been much more of a straightforward task to demonstrate the high inference potential of the site to a sceptical audience, by reference to observations made on site at the time of excavation.

An example of the system at work is the current proposal for a radiocarbon programme for the site. Many small samples of charcoal were encountered in a variety of contexts during the summer 1996 season. Some of these contexts related to zones of arable activity belonging to several phases, probably from the late Bronze Age into at least the Romano-British period. Stratigraphically the episodes are distinct (the matrix, that old Newtonian tool, looks highly convincing), but just how reliable is the provenancing? Using the iconic system it has been possible to rank the samples in terms both of their stratigraphic importance, and their likelihood of providing reliable measurements. This can be done easily by applying ranking points to certain characteristics or combinations of characteristics. Animal or floral disturbance, for example, might score minus 5 points, while a clearly defined horizon might score plus 3 (it sounds like a game of Fantasy Archaeology). As long as the ranking system is explicitly defined, it is possible for a specialist to determine their own rankings by adjusting the weighting of any characteristics. Indeed, weighting the value of characteristics by the use of explicit scoring can lead to many areas of potential interest (the exercise has been tried before, most convincingly by Charles Thomas in his *Christianity in Roman Britain to AD500*) . Evaluative excavations could produce a score for the inference potential of a site which might be weighted by its rarity or culture historical value, but which would at least give developers the chance to scrutinise what must seem to them rather arbitrary opinions on the quality of their site. After all, inference potential is logically a strict function of the relationship between diagnostic and undiagnostic traits (Adams 1991).

More exciting for the excavators, the use of the system at Ingram has provoked the development of new excavation techniques. The soils at Ingram are very difficult because of their homogeneity and because of very poor interface definition. Although landscape sequences here are complex, they have been shown over the last four years to be retrievable, given extreme care. In 1997 a new area was opened on Turf Knowe (Area 13: Adams and Carne 1997) to eliminate the possibility of further structures on the site. After the turf had been hand-stripped a deposit of ploughsoil containing undifferentiated rubble was exposed. In previous seasons such rubble had been cleared down to the first recogniseably intact archaeological horizon, but in this case it was decided to remove all the loose soil and leave every stone in place. This was uncomfortable for experienced archaeologists, emabarrassed at committing the solecism of leaving small stones on little plinths of soil. An archaeologist steeped in the methodology of the single context system might have demanded that every stone be planned as a separate single context on separate sheets of film. Apart from the high probability of graphic error, the cost would have been enormous, and the result no better. The stones were then planned and recorded, and possible structural elements given provisional feature status. Although the entire rubble deposit had clearly been reworked by a plough, there is some evidence (e.g. Lambrick 1980) that actual displacement of material may not always be severe. Using the icons, all possible and probable structural elements within the rubble were evaluated to determine their integrity, for later laboratory analysis.

Subsequently, the rubble was removed, the underlying, partly truncated soil was cleaned, and a rectangular stake-built structure was excavated. No dating evidence was retrieved for the building, and its function can only be marginally inferred from its proximity to other rectangular structures of a probable funerary nature. Crucially, however, later analysis of the records from the rubble layer showed that a circular post-built structure (in form and size a late prehistoric roundhouse) had overlain the rectangular building. Not only that, but careful examination of the characteristics of the post-packing showed that the posts had been deliberately removed. This evidence has now led to a recommendation that all future development control work in that area be constrained by hand excavation of trenches.

The future of the system

The system used by *Archaeological Services* is by no means perfect. Each unit has to find its own best practice, and every excavator has preferences for their ideal system. It has been shown, especially in urban environments, that the sheet as it exists now is inadequate for dealing with specialised contexts such as walls or waterlogged features. The simplest answer is to use the example set by the Toolbox System (Clarke and Stead 1988) and add individual recording tools in an iconic format. These can be simply printed on adhesive paper and attached to the rear of the sheet as required. There is already a set of tools which have been developed and successfully used to excavate inhumations, based on the original query system for Addingham. The more specialised recording tools the better: for example a set of icons for masonry has been developed in association with Adrian Chadwick. Use of the IFPRS means that more intensive techniques can be implemented on site with little extra expense, and with a far greater potential for interrogating a site both during an excavation and in the analytical stage.

Conclusion

The psychologist Dr John Duggan has expressed amazement (pers. comm.) that archaeology has apparentlly never conducted the fundamental experiment that would justify the technique of excavation: that is, how do we know that excavators are recording meaningful data from what they excavate and perceive. Archaeologists know implicitly that this is a complex issue, for the excavator is not only interpreting a fleeting scenario on the spot, but creating many of the characteristics of that scenario in the first place. Such an experiment would be a little tricky to establish, because it would never be possible to replicate conditions exactly. But then, no two rats are the same... In any case, with eye-tracking devices and video recording, it is about time that such an experiment was set up, preferably in collaboration with psychologists.

The author believes that in a decade from now excavators who do not explicitly recognise and record traits of formation processes on site will be regarded by their peers much as excavators are regarded now who do not employ single context recording systems (and they still exist). Perhaps the debate entered into here about the relative merits of subjective or objective methods obscures the message that “what counts is whether the problems are interesting and the tools fruitful, not whether they come from a processual box or a post-processual box” (Bell 1990, 10). Nevertheless, just as archaeologists are concerned about how to match data gathering techniques to our increasing demands for sophisticated information, so psychologists, among others, are keen to understand how humans perceive and process sensory information, and are exploring new techniques for verifying human observations. We should be paying much more attention to what they are saying.

In any case words such as subjective and objective will surely also be regarded as simplistic in the future. The human brain is vastly complex and sophisticated. Too sophisticated, certainly, for us to be happy with the pretence that we only have to measure carefully for us to be scientific. In a decade or so archaeologists will look at our excavation archives and ask themselves how it was that the excavators of the 20th century could fail to record the nature of interfaces, that they could use such loaded terms as good or bad preservation, that they could offer such trite and simplistic categories for complex processes. In the last respect the Durham system will certainly be found wanting, although it is to be hoped that it will be seen as step in a fruitful direction.

Twenty-five years ago Philip Rahtz’s brilliant paper *How likely is likely?* (Rahtz 1975) captured the frustration of an excavator who knew that much of our inference is implicit, who recognised that it was the formation of the archaeological record in its entirety, not just of the site, that we must understand in order to reflect the complexity of the problem that is obvious to all excavators. This author’s plea is for excavators to dare to invent systems that they think will do a better job, which-ever toolbox they have to go to for the right bit.

Archaeological Services - University of Durham: Context record Sheet

SITE CODE:	GRID REF:	AREA CODE:	**CONTEXT #:**
Type:	TST #:	Inc./Filled by:	**Feature #:**
Length x Width (m): X	Depth/Height (m):		Checked by:

Wet Munsell #:	Colour dry/wet:
Composition:	Texture:

Inclusions:

			?						1		H20		An	Fl			
1	2	4	8	16	32	64	128	256	1	2	4	8	16	32	64	128	256

Deposit	Vertical interface	Horizontal interface
Above:	Cuts:	Above:
Below:		Below:
Cut by:	Cut by:	Cut by:
Fill of:	Filled by:	
Same as:	Same as:	Same as:

Qualify process icons:

Physical matrix

Description:

Finds types:

0	P	S	Def	R	u/r		Hz	Fm			
0	1	2	4	8	16	1	2	4	8	16	32

Samples:

Photos C:

Photos B/W:

SMF nos:

Recorded by:

PLANS: SECTIONS:

Date:

Method of excavation:

Computerised:

© Archaeological Services 1995: Version 3

Figure 1: *The front of the ASUD Iconic Formation Processes Recognition System sheet*

Extract from the recording manual

Horizon: Circle the relevant number or numbers to indicate presence, or 'yes'.
The pictures refer to the following concepts:

1	Is the surface of the context a clearly defined horizon? This applies to interfaces as well as deposits - i.e. are the edges clear and easy to determine? For deposits this means the upper surface, while for interfaces it means the interface itself.
2	Is the upper surface poorly defined so that there is a graduation from one entity to the next? This again can also apply to the description of interfaces. Consider the processes involved and correlate if possible with the formation field.
4	Is the upper surface uneven or irregular? This applies to interfaces as well as deposits. Consider how this may have occurred, and correlate with the formation field.
8	Is the upper surface effectively invisible, or are you unsure about where the horizon is? You can use this box to qualify other observations, but you must then fill in the Qualify Process icons box with supporting evidence.
16	Is there evidence of truncation of either a deposit or an interface? (For example, when only the bottom of a post-hole remains). If the truncation was as a result of the archaeological process, this should be noted in the Qualification field, and in the Reliability section.
32	Is the upper surface compacted? If this is so, and especially if there are finds in association with the surface, then a horizontal interface sheet should also be completed.
64	Does the upper surface of the deposit or interface show signs of having been exposed to weathering in the past? Do you need to assign a horizontal interface sheet?
128	For deposits only: is the deposit a laminate? I.e. are there a number of micro-layers which form a discrete deposit - like silting episodes, for example. If yes, then consider whether it is a single episode or a multiple.
256	Is the deposit which you are recording effectively sealed from above? This specifically relates to the integrity of dateable material or environmental samples.

Formation processes

1	Is there evidence that a deposit has been deliberately created in a single episode? This may apply to pit fills and make-up deposits or demolition debris, or to fire episodes. It can also apply to interfaces where there is clear evidence of a single episode, such as a stake hole. If there is doubt the supposition ought to be that it was created over a long period of time, but it is also possible to tick this box and then qualify.
2	Is there evidence that the deposit has accumulated over a long period? Buried soils would be included in this characteristic, as would some laminations, silts and fills of cess or rubbish pits etc.
4	Is there evidence of waterlogging, now or in the past? Qualify.
8	For deposits, is there evidence of a buried soil? Usually this is characterised by a humic layer or the 'cheesecake' like texture of turf.
16	Is there evidence of animal activity e.g. burrowing of rodents or excessive worm holes? Qualify.
32	Is there evidence of floral activity e.g. root holes or dense root fibre penetration? Qualify.
64	Is there evidence that the deposit is wind-derived? E.g. loess. Note that, generally, wind-derived deposits have no large particles such as pebbles.

128 Is there evidence that the deposit has been formed by flowing water, or as a result of standing water. N.B. it is important to determine how silts have been formed. Deposits of iron pan would require this icon to be ticked.

256 Is there evidence that the deposit has been generated as a result of refuse disposal? N.B. to relate this field to the *Finds* field lower on the page. Evidence of nightsoiling in cultivation plots would require this icon to be ticked.

Qualify process icons: Where there is doubt or equivocation concerning icons, they should be accounted for here. Evidence to support inferences should also be entered briefly. This is especially important where there are apparent contradictions recorded as icons. Where animal or floral disturbance has been observed, it should be characterised. This may turn out to be the crucial field in post-excavation, so its completion is of the greatest importance.

Finds: Information about finds should be recorded as follows, but will be subject to specialist editing during post-excavation:

0 No finds retrieved. Small finds should be included when considering this, as well as entered on the reverse of the sheet.

1 *Primary refuse* should be noted here. Primary refuse is that which derives directly from an activity and can be shown to have been deposited where it fell. E.g. flint knapping waste flakes, wood shavings, hearths.

2 *Secondary refuse* is material which has been collected and deliberately discarded. Material from pit and ditch fills and middens usually comes into this category (note cross-referencing to the rubbish code under formation processes). Night-soiling debris on fields also comes into this category, although it may have undergone further movement after deposition. Unless finds can be shown to be otherwise, they will usually be recorded as secondary.

4 *De Facto refuse*. The term applies to material which might have been reused but has been abandoned. Pots lying on floors which may be put back together to form whole vessels come into this category, as do hoards and burial deposits. Coins which appear to have been lost may also be included. Burials always constitute *de facto refuse*, but may have an additional status.

8 *Residual material.* Not to be confused with secondary refuse. Residual material is that which <u>can be shown</u> to be contained in deposits which must derive from a later date. E.g. Roman pottery in a medieval pit, although curation cannot always be ruled out as a process, so this needs to be thought about with care..

16 *Use-wear* and *re-use*. All material should, where possible, be examined for traces of *use-wear* and *reuse* while still on site. Examples include bones with butchery marks, pot sherds reused as gaming counters or clinker, reused timbers, and flints with patination. If ticked, this icon should always be qualified to indicate what evidence has been observed.

Reliability: Information about reliability reflects the quality of the excavation and the archaeology, and should be recorded as follows:

1 *Operational impacts.* In other words, has the excavator either made a mistake such as overcutting or was the archaeology so difficult that it couldn't be retrieved adequately. Qualification must always occur with this icon.

2 *Hz.* Are you concerned about the reliability of the information contained in the Horizon fields? Tick if yes. Qualify.

4 *Fm.* Are you concerned about the reliability of the information contained in the Formation fields? Tick if yes. Qualify.

8 *Time.* Was there too little time to do justice to the quality/complexity of the archaeology? Tick if yes.

16 *Relationships.* Is there a problem with the accuracy of the relationships recorded for this context? If so, qualify.

32 *Ground conditions.* Were the ground conditions too difficult (e.g. too dry, too wet, too cold) for reliable recording? If yes, then qualify.

Acknowledgements
Thanks are due to David Webster, Adrian Chadwick, Chris Cumberpatch, Steve Stead, John Duggan and five 'generations' of Ingram students, for discussions, criticisms and support.

Bibliography

Adams, M. 1987 Formation, Recovery and Inference at Broken K Pueblo: some considerations. *Arch-Form* 1, 4-16

Adams, M. 1991 A logic of archaeological inference. *Journal of Theoretical Archaeology* 2:1-11

Adams, M. 1992a Darwin's worms: a review. *The Field Archaeologist* 18:1-4

Adams, M. 1992b Stratigraphy after Harris: some questions. In Steane, K. (Ed) *Interpretation of stratigraphy: a review of the art,* 13-16 City of Lincoln Archaeology Unit

Adams, M. 1994 *The Ingram and Upper Breamish Valley Landscape Project: Interim Report 1994* ASUD Report 212. University of Durham

Adams, M. 1995 *The Ingram and Upper Breamish Valley Landscape Project: Interim Report 1995* ASUD Report 312. University of Durham

Adams, M. 1999 Beyond the Pale: some thoughts on the later prehistory of the Breamish Valley. In Bevan, B, (Ed) *Northern Exposure: interpretive devolution and the Iron Ages in Britain.* Leicester Archaeology Monographs 4: 111-122

Adams, M. and Brooke, C. 1995 Unmanaging the past: truth,data and the human being. *Norwegian Archaeological Review* Volume 28, No. 2, 93-104

Adams, M. and Carne, P.J. *The Ingram and Upper Breamish Valley Landscape Project: Interim report 1996.* ASUD 420

Adams, M. and Carne, P.J. 1997 *The Ingram and Upper Breamish Valley Landscape Project: Interim Report 1997* ASUD Report 438. University of Durham

Adams, M. and Carne, P.J. 1998 *The Ingram and Upper Breamish Valley Landscape Project: Interim Report 1998* ASUD University of Durham

Adams, M. and Haselgrove, C.C. 1995 Excavations at Port Seton East, East Lothian. *Universities of Durham and Newcastle upon Tyne, Archaeological Reports* 19, 46-51

Adams, M. and Reeve, J. 1989 It's a dirty job but somebody's got to do it: the archaeological context in palaeopathological study. In Roberts, C.A., Lee, F. and Bintliff, J. (eds) *Burial archaeology: current research, mthods and developments. BAR British Series* 211: 267-273

Barham, A. unpub. *Preconceptions when interpreting stratigraphy*

Bell, J.A. 1990 Method in spite of rhetoric. *Journal of Theoretical Archaeology* 1:3-12

Binford, L.R. 1977 Data, relativism and archaeological science. *Man* 22, 391-404

Boddington, A. 1987 Chaos, disturbance and decay in an Anglo-Saxon cemetery. In Boddington, A., Garland, A.N. and Janaway, R.C. (Eds) 1987 *Death, decay and reconstruction: Approaches to archaeology and forensic science,* 27-42

Bronowski, J. 1968 *The ascent of man.* BBC. London

Brown, S.R. 1997 *The history and principles of Q methodology in psychology and the social sciences.* Paper presented at a conference on 'A celebration of the life and work of William Stepehenson. University of Durham 1997

Carne, P.J. and Adams, M. 1995 *Excavations in Darlington Market Place 1994.* ASUD 306. Durham

Chadwick, A. 1998 Archaeology at the edge of chaos: further towards reflexive excavation methodologies. *Assemblage* 3

Clarke, D.L. 1968 Analytical archaeology. 2nd edition. London Clark, P.R. 1988 Studying deposit formation processes on rescue sites: the Toolbox system. *Arch-Form* 3, 6-7

Cumberpatch, C.G. and Blinkhorn, P.W. 1998 The interpretation of artefacts and the tyranny of the field archaeologist. Assemblage 4. www.Shef.ac.uk/^assem/4/1

Gleick, J. 1987 *Chaos: the making of a new science.* New York

Historic Buildings and Monuments Commission (English Heritage) 1991 *The management of archaeological projects.* London

Hodder, I. 1997 'Always momentary, fluid and flexible': towards a reflexive excavation methodology. *Antiquity* 71: 691-700

Lambrick, G. 1980 Effects of modern cultivation equipment on archaeological sites. In Hinchcliffe, J. and Schadla-Hall, R.T. (Eds) *The past under the plough.* London, DOE (Occasional Paper 3), 18-21

Rahtz, P.A. 1975 How likely is likely? *Antiquity* 49:59-61

Schiffer, M.B. 1987 *Formation processes of the archaeological record.* Albuquerque
Shanks, M. and Tilley, C. 1987 *Re-constructing archaeology.* Cambridge

Stephenson, W. 1988 Quantum theory of subjectivity. *Integrative psychiatry* 6:180-187

Thomas, C. 1981 *Christianity in Roman Britain to AD500.* Batsford. London

Webster, D.S. 1998 *GIS and affordance: Llobera's (Non) use of James Gibson.* Unpublished draft copy