

An experimental study of two grave excavation methods: Arbitrary Level Excavation and Stratigraphic Excavation

Laura Helen Evis, Ian Hanson & Paul Nicholas Cheetham

To cite this article: Laura Helen Evis, Ian Hanson & Paul Nicholas Cheetham (2016): An experimental study of two grave excavation methods: Arbitrary Level Excavation and Stratigraphic Excavation, *STAR: Science & Technology of Archaeological Research*, DOI: [10.1080/20548923.2016.1229916](https://doi.org/10.1080/20548923.2016.1229916)

To link to this article: <http://dx.doi.org/10.1080/20548923.2016.1229916>



© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 12 Sep 2016.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

An experimental study of two grave excavation methods: Arbitrary Level Excavation and Stratigraphic Excavation

Laura Helen Evis^{1*} , Ian Hanson², and Paul Nicholas Cheetham³

¹Lecturer in Physical Anthropology, University of Exeter, Archaeology Department, Laver Building, Streatham Campus, North Park Road, Exeter, Devon, UK

²Deputy Director of Forensic Science for Anthropology and Archaeology, International Commission on Missing Persons, Alipasin 45a, Sarajevo 71000, Bosnia and Herzegovina

³Senior Lecturer in Archaeological Science, Bournemouth University, Department of Archaeology, Anthropology and Forensic Science, Talbot Campus, Poole, UK

Abstract The process of archaeological excavation is one of destruction. It normally provides archaeologists with a singular opportunity to recognise, define, extract and record archaeological evidence: the artefacts, features and deposits present in the archaeological record. It is expected that when archaeologists are excavating in a research, commercial or forensic setting the methods that they utilise will ensure a high rate of evidence recognition and recovery. Methods need to be accepted amongst the archaeological and scientific community they are serving and be deemed reliable. For example, in forensic contexts, methods need to conform to scientific and legal criteria so that the evidence retrieved is admissible in a court of law. Two standard methods of grave excavation were examined in this study with the aim of identifying the better approach in terms of evidence recovery. Four archaeologists with a range of experience each excavated two similarly constructed experimental 'single graves' using two different excavation methods. Those tested were the *arbitrary level excavation method* and the *stratigraphic excavation method*. The results from the excavations were used to compare recovery rates for varying forms of evidence placed within the graves. The stratigraphic excavation method resulted in higher rates of recovery for all evidence types, with an average of 71% of evidence being recovered, whereas the arbitrary level excavation method recovered an average of 56%. Neither method recovered all of the evidence. These findings raise questions about the reliability and so suitability of these established approaches to excavation.

Keywords Excavation; Arbitrary; Stratigraphic; Evidence; Stratigraphy; Forensic

Received 12 April 2016; **accepted** 24 August 2016

Background

The process of digging a grave can be considered as a single event of rapid deposition or a 'time capsule' due to the relatively short period of time in which the process is undertaken (Greene 1997; Foxhall 2000). The process of backfilling the grave generally results in a stability in the position of evidence and the human remains present within the grave structure (Hanson 2004).

A grave can be defined as an excavation in the earth for the reception of a corpse (Oxford English Dictionary 2015). As a grave is dug, it 'cuts' the natural and or man-made layers (strata), which are removed and the stratigraphic sequence is disturbed. This process results in the formation of a new surface

(walls and floor) beneath the ground onto which a body or bodies are placed (Hanson 2004). Subsequently, the removed natural/man-made layers are placed back into the grave structure as a 'fill' over the body. Typically, however, these layers become intermixed during their removal and replacement. Differences form in the colour, texture, chemistry, compactness, volume, water retention, odour, organic content and pH level between the disturbed area associated with the grave structure and that of the undisturbed natural/man-made layers through which it was dug (Wolf 1986; Killam 2004). These differences enable the archaeologist to define areas of disturbance allowing for burial locations to be identified and excavated.

*Corresponding author, e-mail address: L.Evis@exeter.ac.uk

In normal archaeological fieldwork, the process of excavating a grave is perceived as a simple one. During excavation the grave cut is defined and the grave fill is often found to be a single stratigraphic deposit, and is removed as such, whilst the body is viewed as an artefact (Browne 1975; Hunter 1994). In practice, the stratigraphy of graves can be much more complex, for example in cemetery contexts where there are multiple internments over time. In forensic contexts, a grave is considered to potentially contain multiple layers that can be recognised including those of organic decomposition and additives such as lime that may have been used to assist in a grave concealment process (Hunter 1994; Congram 2008). If the original grave fill has later come to be disturbed, for example by the perpetrator or animal activity, the grave structure may then contain several different cuts and fills (Hochrein 2002).

Normal archaeological excavation methods have had to be adapted in the light of the potentially complex nature of recent burials and their forensic investigation (Hunter 1994). This adaptation is largely characterised by processes to establish forensic relevance, limit contamination, record stratigraphy using spits and sections across the grave, as well as the retention of grave fills for subsequent detailed analysis. However, as with field archaeology generally, the methods utilised and published by forensic archaeologists/anthropologists vary extensively. They have evolved to their current state according to the archaeological practices advocated by practitioners and professional bodies in their country of origin, and the inherited traditions present in each. Consequently, different excavation methods and recording systems are used by different archaeological practitioners in accordance with their individual preferences, which are largely formed by the site types from which an archaeological practitioner has gained their academic training and experience (Carver 2009; Carver 2011:107). Two principal methods, the *arbitrary level excavation* method and the *stratigraphic excavation* method have developed through different traditions and archaeological needs.

Arbitrary level excavation

As part of their academic training physical anthropologists and archaeologists may receive training in archaeological field schools, many of which emphasise excavation by arbitrary levels as a standard approach. This method is commonly utilised in test pitting in professional archaeological assessments, contributing to the wide scale adoption of the arbitrary level excavation method in forensic casework. Practitioners using this method have published technical papers regarding the forensic application of archaeological techniques, and as a consequence, the arbitrary level excavation method has come to be regarded as a standard excavation method for forensic investigations (Ramey-Burns 1996; Crist 2001; Komar and Buikstra 2008).

During the arbitrary level excavation of a grave, soil is removed in a succession of predetermined levels, usually 0.05 m, 0.10 m, or 0.20 m in depth (Hester 1997:88), over an arbitrary but carefully measured area, usually determined by the perceived size of the grave at surface level. As evidence is identified the earth 'matrix' that surrounds it is removed leaving each item upon a soil 'pedestal'. These items are measured in situ and only removed when they are deemed to be hindering the progress of the excavation (Joukowsky 1980; Brooks and Brooks 1984; Ramey-Burns 1996; Tuller and Đuric 2006; Connor 2007). During this process, soils that comprise the deposits backfilling the grave as well as the surrounding natural/man-made strata through which the grave was originally dug are removed in spits across the defined area of excavation. In order to provide access to the burial, trenches are often dug around the remains resulting in the removal of the grave walls (Joukowsky 1980; United Nations 1991; Godwin 2001:9). Some practitioners advocate against the removal of the grave walls however, as these surfaces may be of assistance when interpreting the method by which the grave was constructed, and assist investigators in establishing links between the crime scene and the perpetrator(s) (Powell et al. 1997; Hochrein 2002; Dupras et al. 2006; Connor 2007).

The arbitrary level excavation method has several perceived advantages, including: spatial and depth control of soil removal and artefact recovery; easier access to the remains and artefacts from different angles; dynamic photographs can be taken of both the human remains and artefacts; it assists with potential water drainage issues that can damage the integrity of the grave structure; and it limits the time spent standing on a grave structure of limited size that could damage the human remains and artefacts (Spennemann and Franke 1995; Pickering and Bachman 1997; Godwin 2001; Hochrein 2002; Tuller and Đuric 2006). Notionally less archaeological skill and experience are required to utilise this method, as spits can be easily measured and levelled to accurate standard depths.

However, there are inherent problems with this method, including: the method destroys and ignores stratigraphic interfaces and layers present within the grave; it introduces artificial divisions of deposits and evidence which can result in evidence retrieved during the process of an excavation having no known stratigraphic origin; it results in the mixing of strata and artefacts from the grave structure (fills and cuts) and natural strata through which the grave was dug potentially leading to contamination of soils and artefacts that may pre or post-date the grave; the grave walls can only be recorded in plan at the interface of each arbitrary level (if distinguishable from the natural strata) which will not always allow for the accurate recording of the grave cut including tool marks; and pedestalled artefacts may be moved during excavation (Harris 1979, 1989,

2002; Hanson 2004; Hunter and Cox 2005; Komar and Buikstra 2008).

Despite these weaknesses, the arbitrary level excavation method continues to have advocates for its application; it has been argued that this is largely due to the fact that, normally, graves lack complex stratigraphy and are usually comprised of a singular fill, and therefore the application of arbitrary units is justifiable. The primary emphasis when utilising this method is often upon the recovery of artefacts and human remains, rather than understanding the entirety of the grave formation process. Arbitrary level excavation provides the easiest and most efficient method for meeting this objective (Pickering and Bachman 1997; Haglund et al. 2001).

However, it may be necessary to demonstrate as complete a stratigraphic record as possible has been recognised and excavated, and that the evidence of that stratigraphic record has not been lost, but was recovered and documented. It is a normal requirement that excavation should be undertaken to a standard that allows re-interpretation from the documentation. Archaeologists may therefore need to demonstrate they have recorded the basis to accurately interpret the stratigraphic record, record the stratigraphic sequence, and justify the reconstruction of the sequence of human and taphonomic events that occurred at the site under investigation (Harris 1989, 2002; Hanson 2004). It has been argued that the best way that this can be achieved is through the use of the stratigraphic excavation method (Barker 1987, Harris 1989, 2002; Hochrein 2002; Hanson 2004).

Stratigraphic excavation

When using this method, separate archaeological stratigraphic contexts are identified and excavated individually in sequence, and recorded as individual stratigraphic phenomena. The entire grave is viewed as an archaeological feature. Thus the fills and interfaces are normally revealed and recorded in their entirety and grave walls may be exposed and maintained throughout the entire excavation process. This allows for the retention of tool marks and geotaphonomic evidence present on the surfaces of the grave walls and grave floor (Hochrein 2002).

There are several perceived advantages to stratigraphic excavation including: three dimensional recognition, assessment and recording of each stratigraphic context; revealing of interfaces between deposits; chronological recovery of evidence by context; spatial and depth control of soil removal and artefact recovery; prevention of contamination between stratigraphic contexts; dynamic photographs can be taken of both the human remains and artefacts reflecting their chronological deposition; and removal of deposits that records the sequence of deposition to aid in the reconstruction of events.

The main problems with this method are: without tents and other precautions water can collect in the grave; that excavation in limited spaces and at depth

can limit access to the human remains (Tuller and Đuric 2006); difficulties in recognising individual stratigraphic contexts, especially interfaces; that the method is more complicated to perform than other methods; and that the method may be perceived to slow down excavation.

In normal archaeological excavation, differences in interpretation or implications of mistakes made during the excavation and interpretation of archaeological sites are not seen as inherently problematical. However, differences in interpretation, misinterpretations or destruction or loss of evidence during the excavation process in forensic contexts have potentially greater ramifications. The results from such work have significant legal, political, social and media impact. Loss of evidence may impact investigations and prosecutions, and in some countries, for example Iraq, there are legal penalties (fines and imprisonment) for evidence loss (Crist 2001; Law on the Protection of Mass Graves 2006). It is therefore prudent that excavation methods are assessed and tested to determine their suitability. Establishing whether there may be error rates, variation in results and impacts on interpretation depending on methods used is a sensible scientific aim. Given that each archaeological site is unique, how excavation methods can be compared raises questions about how to approach experiments to assess this. An experiment was designed so that the arbitrary level excavation method and stratigraphic excavation method could be tested in a controlled environment. This would compare evidence recognition, recording and recovery rates for typical evidence forms present within a grave site when excavated by participating archaeologists. The timeframes of the experiment (concerning the creation and excavation of the artificial features) matched that seen in forensic casework, where there is often a limited time between burial and recovery.

Experimental design

In order to allow for the objective comparison of the stratigraphic excavation method and arbitrary level excavation method it was decided that artificial features with similar properties to single graves would be utilised. They were designed to be as identical as possible to each other in regards to their location and properties: shape, size, archaeological contexts and evidence. The aim was to minimise the number of variables that could affect evidence recovery, and standardize the structure and content to ensure that each method could be directly compared. During this experimental study evidence was defined as: artefacts, tool marks, and stratigraphic contexts (deposits/fills, cuts/interfaces).

The 'graves' were created using a mechanical digger. This was deemed justifiable as mechanical diggers are commonly used to dig graves (Hunter and Cox 2005). Through using a mechanical digger the researchers were able to impose standard

dimensions and also distinctive tool marks on the walls and base of the graves, which, if identified, would assist the archaeologists in their interpretation of how the grave was constructed. Each grave measured 1.20 m in length, 0.75 m in width, and 0.85 m in depth. Approximately 2.0 m was left between each experimental grave to ensure that an adequate working space was left for the excavations to be undertaken.

The experimental graves did not contain any form of skeletal remains as this experimental study was not concerned with the osteological recovery potential of the two excavation methods, something explored by Tuller and Đuric (2006). Morse et al. (1976a; 1976b) discuss how they created 'graves' with no skeletal remains for the purposes of training investigators in forensic archaeological excavation procedures. Therefore, the researchers classified these cut features as graves despite the absence of skeletal remains, but with the expectation of participants that remains were present.

The artefacts (Figure 1.0; 2.0) that were included in the graves were chosen to represent items typically found in clandestine burials. In addition, it was determined that these items would preserve during the short time between their burial and subsequent excavation (Janaway 1996; Janaway 2002). These items were also common, easily identifiable items, and thus would be recognisable to participants. These items also varied in size, composition and shape enabling the researchers to determine if excavation (by either method) had a tendency to recover artefacts of a certain size, composition or shape. Several soil fills were used to back fill each grave cut. A secondary cut was made into these fills, which was itself filled. Artefacts were placed within these fills and on interfaces (Figure 1.0; 2.0). The depth and distribution of each stratigraphic context was matched to be the same in each grave. All artefacts were placed in the same location in each context in all graves and the locations recorded in three dimensions. Moreover, according to scholars such as Hanson (2004) and Hunter and Cox (2005), the arbitrary level excavation

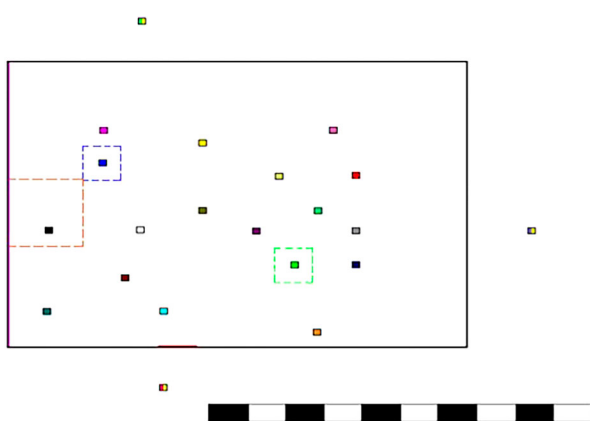


Figure 1.0 Illustrates the position of evidence within the grave.

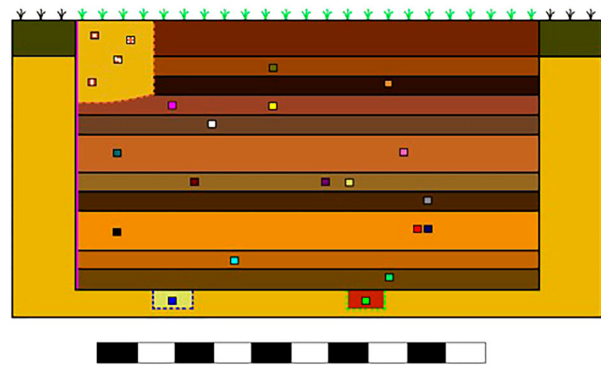


Figure 2.0 Illustrates the position of evidence within the grave and the structure of the grave.

method can result in the mixing of artefacts from a grave fill with those present within the natural undisturbed strata through which the grave was dug, thus resulting in the collection of evidence unrelated to the grave creation events. The stratigraphic excavation method can also lead to over-excavation of contexts as the excavator seeks to define interfaces and the edges of deposits. In light of these observations, the researchers created incisions into the natural undisturbed strata 0.15 m beyond the edge of the grave cut into which a key, marble and coin were placed. Such items are those that could easily be lost at the site prior to or after the graves creation. Through the inclusion of such evidence in the experiment, the researchers could assess if excavation would result in extraneous evidence being retrieved.

In all, eleven distinct horizontal deposits were added to each grave. Although the presence of multiple perfectly horizontal deposits are, as Praetzelis (1993:18) states, the "exception rather than the rule" in archaeological sites, following this procedure made the exact replication of each grave and matched positioning of the contents achievable, accurate and efficient. One potential effect of horizontally placed deposits is that the excavated arbitrary 0.10 m levels could coincide with the horizontal deposit interfaces within the grave fills. This may favour recognition of evidence during arbitrary level excavation. The stratigraphic sequence was made more realistic and less uniform by varying the depth of deposits from 0.05 m and 0.10 m. Moreover, the inclusion of the internal feature and associated fill cutting the primary fills of the grave, and two additional cut features and associated fills in the floor of the graves allowed both methods to be compared through the potential to reveal a number of vertical and horizontal interfaces (Figure 2.0; 3.0).

Additionally, all graves were left exposed to the elements for seven days. This was intended to produce the typical geotaphonomic phenomenon of surface cracking (Figure 4.0). In experiments conducted by Hochrein (2002: 55), it was noted that such phenomenon can be recovered during excavation and can be indicative of a grave feature being

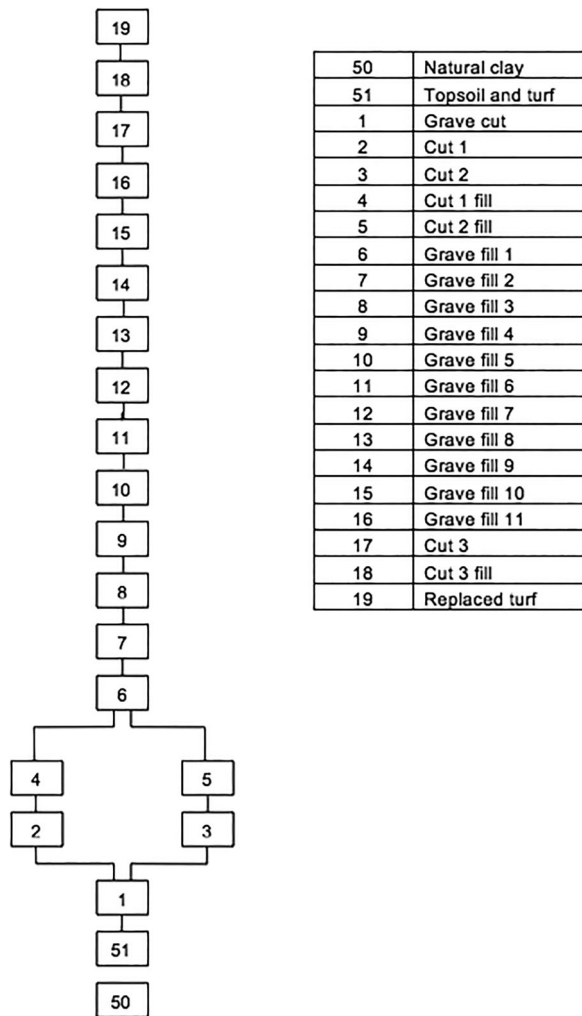


Figure 3.0 Illustrates the stratigraphic relationships of the stratigraphic contexts contained within the grave.

prepared in advance of a homicide event; thus providing a sign of premeditation. To further this concept of a pre-prepared grave, leaf litter from the surrounding area was placed into the bottom of the grave. As Hunter and Cox (2005: 109) note, the presence of vegetation in the bottom of graves can be indicative of a



Figure 4.0 Illustrates the surface cracking evident on the grave walls.

grave that has been left open for a time before infilling. The inclusion of this vegetation layer disguised the ‘true’ grave floor, providing a qualitative test for the archaeologists during the excavation experiment, to see if they excavated the grave until the floor of the grave or ‘sterile’ deposits were reached, as recommended in forensic archaeological excavation literature (Hunter and Cox 2005). Each grave was covered with loose soil and turf so that visually the general outline of each grave was not visible at surface level. The graves were set up in natural stratigraphy of leached grey and orange sand with iron panning, over gravel layers. The fills used in the graves were formed from the material removed during the machine excavation, except for the layer of leaf litter.

Other factors taken into consideration were that each archaeologist would be excavating two replica graves each using different methods, and that multiple archaeologists would be excavating their graves at the same time, with the potential to overlook or communicate with neighbouring excavators. To prevent the former factor from being an issue, the graves were arranged in sets of two, which were 180° mirror images of one another. This was so excavators would not recognise the properties of the second grave they excavated compared to the first grave. In addition, at no point were the archaeologists informed that the graves were identical in terms of dimensions and content. Moreover, from the findings of previous researchers such as Harris (1979; 1989; 2002), Hanson (2004), Tuller and Đuric (2006), and Komar and Buikstra (2008), it was evident that the arbitrary level excavation method could be expected to intercut the different stratigraphic contexts contained within the graves and destroy certain forms of evidence, including: the grave walls and tool marks. Therefore, each archaeologist was told to use the arbitrary level excavation method for their first grave excavation. Although this represents a clear bias in the organisation of the experiment it was deemed justifiable as it would assist in reducing the overall impact of participants potentially recognising similarities between their graves. To combat the latter factor, forensic tents were placed over the graves to limit views, whilst they were excavated and tarpaulins were placed over the graves when the site was left. The participants also agreed not to talk with one another until the experiment had finished.

Each of the participants were self-selecting volunteers, but were required to have had varying experience in the excavation of grave features. Archaeologist 1 had gained seven days of archaeological excavation experience and had excavated one grave previously. Archaeologist 2 had gained three months of archaeological excavation experience and had excavated two graves previously. Archaeologist 3 had obtained two and a half years of archaeological excavation experience and had excavated five graves previously. And Archaeologist 4 had six years

archaeological excavation experience, and had excavated over 100 graves.

Excavation and recording equipment

Participants were able to select excavation and recording equipment from the following: mattock, shovel, digging spade, buckets, trowel, hand shovel, sieve, tape measures, ranging poles, scales, line level, plumb bob, string, photographic board, cameras, drawing board and permatrace.

For the arbitrary level excavation method, the archaeologists were provided with a recording pack containing spit-level forms, unit-level forms, an artefact register, a photographic register, a drawing register, and a human remains recording form. Whereas, for the stratigraphic excavation method, the archaeologist's recording pack contained context recording forms, an artefact register, a photographic register, a drawing register, and a human remains recording form. Observation sheets were provided to the excavators so they could describe the process they were undertaking.

Excavation procedure

Method guidance documents were provided for the arbitrary level excavation method and were adapted from the excavation guidelines outlined in Ramey-Burns (1996; 2006) and Connor (2007) (see Appendix 1). The use of Ramey-Burns' method guidelines was deemed appropriate as she had also contributed to the formation of the United Nations excavation guidelines (1991), which have been used globally during international investigations of human rights violations. The participants were briefed on the method order to employ and that they were excavating graves. Following provision of the aforementioned guidance and the recording forms, the archaeologists defined the outline of the grave cut; they then delineated an area larger than the grave - 3.0 m in length by 2.0 m in width using pegs and string. Each archaeologist proceeded to remove the overlying turf and first 0.10 m spit using available tools. Once the first 0.10 m spit was removed the archaeologists continued to excavate in arbitrary 0.10 m levels. When an artefact was identified its location was recorded in three dimensions and spit-level noted, it was then left upon a soil pedestal. All evidence and associated pedestals were left in place until the individual excavator decided that it was hindering the progress of their excavation. The evidence was then removed and the pedestal excavated. All soil removed during the excavation of each spit was kept separate from other spits and was sieved. The final spit, 0.80 m to 0.90 m took the archaeologists to the depth of sterile soil (see Figure 5.0).

The method guidance documents provided for the stratigraphic excavation method were adapted from the excavation guidelines outlined by the Museum of London Archaeology Service (1994), Hanson (2004), and Hunter and Cox (2005) (see Appendix 2). Following provision of the aforementioned guidance and the



Figure 5.0 Illustrates the excavation of a grave using the arbitrary level excavation method.

recording forms, the archaeologists defined the outline of the grave cut. The archaeologists then excavated each fill/deposit they observed within the grave and maintained the boundaries of any interfaces identified. Each of the interfaces and fills/deposits recognised were treated as unique (contexts) and any fills/deposits were stored and sieved separately. When an artefact was identified, its three dimensional location was recorded and context noted. The grave walls were kept intact throughout the entire excavation process (see Figure 6.0).

Throughout the experimental excavations, the archaeologists were observed and their actions documented using voice notes, written notes and photographs. The researchers ensured that they did not



Figure 6.0 Illustrates the excavation of a grave using the stratigraphic excavation method.

communicate with the archaeologists during experimental testing so as to minimise any potential biases.

Results and Discussion

The results presented in this paper focus on the recovery of archaeological evidence. Results relating to the recording and interpretation of archaeological evidence will be reported elsewhere.

Artefacts

Each of the four participants excavated one grave using the arbitrary level excavation method and then another using the stratigraphic excavation method. No participants recognised that the graves had identical properties, or were a 180° mirror image of each other. All participants used the tools and materials available. They did not communicate with each other. They provided feedback on their excavation, methods employed and issues encountered by completing observation sheets as the excavations progressed.

Using the arbitrary level excavation method an average of 64% of artefacts were recovered (Table 1.0). The rate of retrieval varied between 55–77% amongst the archaeologists (Table 1.0). Artefacts were found both in the locations in which they had been placed and out of situ (where items were moved during excavation). The amount of artefacts found out of situ varied from 18–54% (Table 1.0). There was a distinct correlation between the time that an archaeologist spent excavating and the amount of artefacts that were found out of situ, with the more time spent excavating leading to more artefacts being found in situ. Through observing the archaeologists whilst they were using the arbitrary level excavation method it was apparent that the recovery of artefacts out of situ can be attributed, in part, to the method itself; when the archaeologists were trenching around the suspect grave cut area in order to create an access trench using a mattock the archaeologists inadvertently removed the edge of the grave fill, where the definition between the natural undisturbed strata and grave fill was less distinct, resulting in some artefacts situated near the edge of the grave cut being knocked out of situ and recovered during sieving. Despite finding artefacts out of situ, the archaeologists were able to reassociate artefacts with the spit from which they had originated and determine their relative depositional sequence. However, all archaeologists failed to identify all of the contexts within the grave structure, and subsequently associated some of the recovered artefacts with the incorrect contexts. The extent to which they were incorrect varied in accordance with the number of contexts correctly identified, with the accuracy of the interpretation of the depositional sequence of artefacts placed into the grave averaging at 51%, with a variance rate of 4% (Table 1.0).

An average of 72% of the placed artefacts were recovered using the stratigraphic excavation method, with the total artefact retrieval rate varying between 59–82% amongst the archaeologists (Table 1.0). Each of the archaeologists identified artefacts both in the locations in which they had been placed and also out of situ (where items were moved during excavation). Artefacts identified out of situ were recovered by sieving individual contexts. The amount of artefacts found out of situ varied from 0–46% (Table 1.0). As found with the arbitrary level method, there was a distinct correlation between the time that an archaeologist spent excavating and the amount of artefacts that were found out of situ. Despite finding artefacts out of situ, due to the archaeologists using the stratigraphic excavation method archaeologists were able to reassociate the artefacts that they had recovered in the sieve with the context (deposit/fill/interface/cut) from which the artefacts had originated. Thus they were able to place these items within the stratigraphic sequence of the grave and determine their relative depositional chronology. However, all of the archaeologists failed to define all of the contexts within the grave structure. They subsequently associated some of the recovered artefacts with the incorrect contexts, making their reconstruction of the stratigraphic sequence and overall interpretation of the artefacts deposition sequence incorrect. However, the extent to which their reconstructions were incorrect varied in accordance with the number of contexts correctly identified, with the accuracy of the interpretation of the depositional sequence of the artefacts placed into the grave averaging at 71%, with a variance rate of 38% (Table 1.0).

Extraneous artefacts

As stated earlier, the arbitrary level excavation method could result in the mixing of artefacts from the grave fill with those present in the natural undisturbed strata through which the grave was dug, leading to the inclusion of artefacts unrelated to the grave creation event. The inclusion of a marble, key and coin outside the grave boundary, within the natural undisturbed strata tested this supposition. Whilst utilising the arbitrary level excavation method two archaeologists recovered extraneous artefacts – marbles and coins (Table 1.0). The close proximity of these items to the boundary of the grave cut and subsequent pedestal of these items resulted in these archaeologists being unable to distinguish these items as unrelated to the grave structure, and therefore, mistakenly categorised these items as artefacts related to the grave. The other two archaeologists did excavate the areas containing the extraneous artefacts, but failed to recognise or locate any of the items. Whilst utilising the stratigraphic excavation method, one archaeologist identified an extraneous artefact (Table 1.0). The recovery of the key occurred whilst this archaeologist was attempting to define the boundaries

Table 1.0 Illustrates the overall results of the experimental research.

	AF recovered	AF recovered in situ	AF recovered out of situ	Extraneous AF	SC recovered	Tool marks recovered	Time taken (hrs)	Total evidence (AF+SC+TM) recovery
SE: Archaeologist 1	82%	83%	17%	0%	62%	100%	17	73%
SE: Archaeologist 2	59%	54%	46%	0%	52%	50%	11	56%
SE: Archaeologist 3	68%	100%	0%	0%	81%	50%	9	73%
SE: Archaeologist 4	77%	47%	53%	33%	90%	50%	8	82%
<i>SE Method Average:</i>	72%	71%	29%	8%	71%	62.5%	11¼	71%
ALE: Archaeologist 1	77%	82%	18%	67%	52%	0%	31	62%
ALE: Archaeologist 2	55%	50%	50%	0%	48%	0%	16	49%
ALE: Archaeologist 3	64%	64%	36%	67%	52%	0%	23	56%
ALE: Archaeologist 4	59%	46%	54%	0%	52%	50%	8	56%
<i>ALE Method Average:</i>	64%	62.5%	37.5%	33%	51%	12.5%	19½	56%

Key information:

Archaeologist 1: 7 days of archaeological experience,
 Archaeologist 2: 3 months of archaeological experience
 Archaeologist 3: 2.5 years of archaeological experience
 Archaeologist 4: 6 years of archaeological experience
 SE = Stratigraphic excavation method
 ALE = Arbitrary level excavation method
 AF = Artefacts
 SC = Stratigraphic contexts
 TM = Tool marks

*The total evidence recovery is the sum of the artefacts (in and out of situ), stratigraphic contexts and tool marks recovered expressed as a percentage of the total of the three classes of evidence.

of the grave cut, and mistakenly overcut the grave edge, leading to the recovery of the key.

Stratigraphy

Through following the arbitrary level method of excavation each archaeologist proceeded to remove a 2.0 m × 2.0 m area that included the grave structure and surrounding natural strata in a series of 0.10 m spits. Through excavating using this method, an average of 51% of the stratigraphic contexts were correctly identified (Table 1.0). There was little variance in the number of stratigraphic contexts correctly identified using this method, with the results ranging from 48–52% (Table 1.0). All of the archaeologists were able to identify the grave cut as the grave fill was distinct from the natural undisturbed strata, and were able to measure its dimensions all the way to the base of the grave, as all of the archaeologists' spits coincided with the grave floor. The archaeologists could map the grave cut's dimensions, in plan form only, as the method itself had destroyed the grave structure as spits were removed. All of the archaeologists failed to identify and define the presence of secondary cuts within the grave structure. This is due to the method itself, as the approach did not require archaeologists to look for or maintain evident interfaces within the grave structure. By not maintaining the limits of interfaces, the archaeologists found it difficult to identify and define the stratigraphic contexts present. Ultimately, this resulted in the archaeologists being unable to define the chronology of activity within the grave structure; the artefacts that were placed into the secondary cuts becoming intermixed and grouped with the artefacts retrieved from the primary grave fills. The failure of all of the archaeologists to identify all of the primary grave fills was the result of the method. Eight of these fills were 0.05 m in depth, thus as the archaeologists excavated using their 0.10 m spits they inadvertently excavated two fills within one spit, resulting in the combining and intermixing of the fills and the artefacts contained within them.

Through following the stratigraphic excavation method each archaeologist proceeded to remove each individual deposit/fill, defined by differences in texture (the size of the soil particles), composition (types of organic and inorganic matter), volume, compactness and colouration. They did so in the reverse order in which they were deposited, from the latest to the earliest. This method approach enabled the archaeologists to define the interfaces/cuts present. This meant that any 'cuts' identified by the archaeologists during the excavations were defined as a unique event (context), and any fills/deposits contained within them were excavated separately. This allowed the archaeologists to document different phases of activity present within the grave structure, and in turn, separate any of the artefacts recovered into the different stratigraphic phases of deposition present within the grave structure. An average of 71% of the stratigraphic

contexts (deposits/fills/interfaces/cuts) were correctly identified whilst using the stratigraphic excavation method (Table 1.0). However, the number of stratigraphic contexts correctly identified varied significantly between archaeologists from 52–90% (Table 1.0). One archaeologist failed to identify the secondary cut and associated fill at the top of the grave, and three archaeologists did not identify the secondary cuts found at the base of the grave and their associated fills. One archaeologist correctly identified all of the primary fills contained in the grave structure. However, as one archaeologist was able to identify all of the primary grave fills present and another archaeologist was able to define all of the secondary cuts and associated fills within the grave structure it demonstrates it was possible to do so. It suggests the failure by some of the archaeologists to identify and define all of the stratigraphic contexts present in the grave may not have been due to the method itself but other factors such as excavation experience, ability and the observation skills of the individual archaeologist.

Tool marks

The arbitrary level excavation method recovered an average of 12.5% of tool marks present within the grave (Table 1.0). Only one archaeologist identified the presence of a machine bucket tool mark because the archaeologist's final spit coincided with the grave floor, which maintained the imprint of the bucket teeth. As a result, the archaeologist was able to determine that the grave was created using a mechanical digger. All of the other archaeologists failed to identify the presence of any tool marks. This can be attributed to the method itself as the arbitrary level excavation method followed by the archaeologists destroyed the grave walls and tool marks while developing access to the grave, leading to three of the archaeologists being unable to determine how the grave was constructed.

The stratigraphic excavation method recovered an average of 62.5% of the tool marks present within the grave (Table 1.0). All of the archaeologists were able to identify the presence of machine bucket tool marks. They were therefore able to discern how the grave was constructed. Only one archaeologist identified the mattock mark along the grave wall. The failure of three of the archaeologists to identify the mattock mark is not accountable to the method itself, but the observation skills of the individual excavator, as by utilising this method the grave walls were maintained and therefore all tool marks were potentially recoverable.

Time

There was a significant difference in the number of hours it took to complete the excavation of the graves using the two methods. Whilst excavating using the stratigraphic excavation method the archaeologists took an average 11¼ hours to complete the

excavation, although the time spent excavating varied between 8–17 hours amongst the archaeologists (Table 1.0). In comparison, whilst excavating using the arbitrary level excavation method, the archaeologists took an average of 19½ hours to complete the excavation, but the time spent excavating varied between 8–31 hours amongst the archaeologists (Table 1.0). The difference in the length of time that it took for the archaeologists to complete the excavation is largely due to the requirement of the arbitrary level excavation method to remove both the natural undisturbed strata as well as the stratigraphic contexts contained within the grave itself, resulting in over three times the volume of soil (and more compact soil) needing to be removed in order to complete the excavation. Approximately 2.8m³ of soil being extracted and sieved using the arbitrary level excavation method and 0.8m³ using the stratigraphic excavation method. This accounts for the greater length of time it took for the archaeologists to complete the excavation of the grave using the arbitrary level excavation method. In addition, the need to remove three times the volume of material to excavate the same feature may also compromise recovery rates as a result of increased fatigue.

Experience

In regards to experience, the results indicate that higher levels of experience have a positive impact on overall performance and evidence recovery (Table 1.0). Only Archaeologist 1, who had the least experience, did not follow this trend. This result can be explained by the fact that this participant spent between 6–9 hours longer than the other participants excavating using the stratigraphic excavation method, and 8–23 hours longer than the other participants using the arbitrary level excavation method (Table 1.0). Through using this extra time the participant was able to successfully identify more evidence than one might have expected, given their lack of experience. These findings highlight that time as well as experience are key variables in improving overall performance and evidence recovery in archaeological investigations; the greater the length of time spent excavating and the more archaeological experience gained, the better the overall evidence recovery process will be. This has important implications for forensic investigations where pressure is placed on forensic archaeologists to finish their investigative work as quickly as possible. These results show that such time constraints could reduce the volume of evidence recovered and thus the reliability of the investigative team's findings.

Conclusion and Recommendations

The results gained from this comparative excavation experiment indicate that the stratigraphic excavation method was the most productive in terms of total evidence recovery; with all participants achieving

consistently better recovery rates of relevant artefacts, stratigraphic contexts and tool marks. While both methods recovered the majority of artefacts, participants using the stratigraphic method were consistently more successful at identifying the stratigraphic contexts, especially the interfaces and surfaces. Moreover, when using the arbitrary level method, the participants consistently destroyed both the vertical and horizontal interfaces present. The stratigraphic excavation method also proved to be a faster method of excavation, as the arbitrary level excavation method required a greater volume of soil, and consolidated undisturbed deposits to be removed.

When using the stratigraphic excavation approach, the archaeologists were more able to determine the method by which the grave was created. Moreover, due to the retention of the grave walls during excavation, the archaeologists were able to identify the surface cracks between the grave walls and fills, as well as define the layer of vegetation at the bottom of the grave. They were therefore able to suggest that the grave may have been left open prior to backfilling. The arbitrary level excavation method also allowed for the recovery of the vegetation layer, but due to the destruction of the grave walls, the archaeologists were unable to identify the surface cracks. Consequently they could also suggest that the graves had been left open prior to backfilling, but with less certainty than with the stratigraphic excavation method.

The arbitrary level excavation method also resulted in four items of extraneous evidence being recovered. This has implications for the dating of contexts and features. In forensic settings, if items such as these were recovered and thought to be related to the criminal events and the grave structure when they were not, it could result in a considerable waste of investigative time and resources, misdating of the grave feature, the incorrect identification of potential murder weapons, and false leads to identify perpetrators.

On the basis of the results of this limited experimental study, the stratigraphic excavation method is more appropriate for the excavation of single graves, due to its ability to consistently recover a greater percentage of evidence types than the arbitrary level excavation method regardless of experience or skill level. While the arbitrary level excavation method is often deemed easier to undertake, and the stratigraphic excavation method is perceived as more complex to employ, all of the archaeologists consistently achieved a better rate of success in recovering all evidence using the stratigraphic excavation method, despite variation in their experience levels.

This small-scale experiment was designed primarily to compare excavation methods applied to the same stratigraphic sequence, with the same tools and background information available to excavators. The experiment did not allow for variation in method on each grave. In this way the normal flexibility of approach to excavation archaeologists may apply was limited, this was deliberate as the aim was to test a method

as a standard approach. The experiment did not have enough participants to assess in depth or statistically the impact of experience and skill of excavators on the implementation of methods and rate of evidence recovery.

However, the fact that neither method was able to recover all evidence contained within the grave(s) in this experiment is of interest, considering the excavators were provided with the tools that would allow all evidence to be found. Variation in how excavation methods reveal the archaeological record and how those methods are employed should be of concern for all archaeologists. Given the usage of excavation methods in criminal casework, it is therefore of importance that researchers investigate why there is variation and how evidence recovery rates can be improved. Similar research is being undertaken in a range of scientific disciplines that are applied to legal work (NAS 2009).

It is evident that there is a lack of standardisation in regards to the application of traditional archaeological excavation methods even in forensic archaeology (see for example Groen et al. 2015). This is largely a reflection of the lack of standardised practices in commercial archaeological and research-led fieldwork practiced globally; a variety of favoured excavation methods are employed regionally around the world (see for example Carver et al. 2015). These methods have been directly adopted into forensic fieldwork. Where the stratigraphic excavation method and arbitrary level excavation method are actively used, they are often used exclusively, rather than as part of a range of methods that best suit the nature of the site under investigation. Any method used during the course of a forensic investigation may be required to be subjected to empirical testing in order to ensure that it is reliable and therefore admissible (Daubert Standards 1993; Rule 702 2000; Hunter and Cox 2005), and it should be presumed that this will be the case. Nevertheless, little research has been conducted to experimentally test archaeological methods and so establish such reliability.

The assessment in this small study of these two common archaeological excavation methods should be viewed as a pilot study to test the applicability of this experimental approach, and it has provided useful results to use to develop further studies and stimulate discussion. While it is important for archaeology as a discipline to consider assessment of excavation methods, and indeed there is an ethical impetus to undertake the best possible practice (see Harris 2006), it is in stringent legal contexts that a lack of empirical testing of methods can impact whether evidence is accepted in a court of law. In order for forensic archaeology to continue to develop as a discipline, it is recommended researchers continue to experimentally test archaeological excavation methods as well as recording systems to ensure that they are suitable for use in forensic practice. There are clear consequences to not doing so.

Acknowledgements

This research was undertaken at Bournemouth University in 2009. The authors would like to thank the landowner, William Bond, for permission to use the site on the East Holme Estate, Dorset. The authors would also like to thank the individuals who participated in the study.

Funding

The first author privately funded this work.

Disclosure statement

No financial benefits or interests have arisen from the direct application of this research.

Legislation

- Daubert Standards (1993)
- Rule 702 (2000)
- Law on the protection of mass graves: Law Five (2006)

ORCID

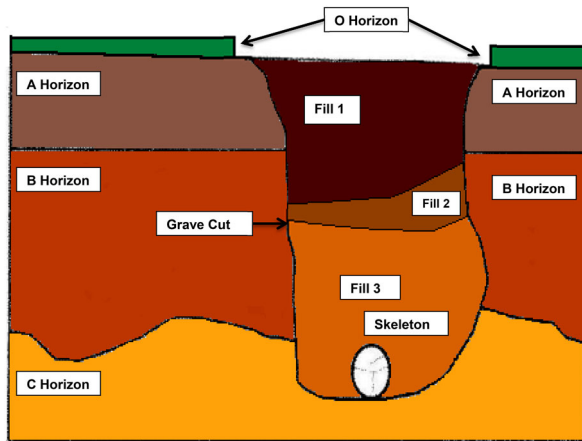
Laura Helen Evis  <http://orcid.org/0000-0002-8493-5166>

References

- Barker, P. (1987). *Techniques of archaeological excavation*. London: Batsford.
- Brooks, S.T. & Brooks, R.H. (1984). *Problems of burial exhumation, historical and forensic aspects*. In T.A. Rathbun., & J.E. Buikstra, *Human Identification Studies in Forensic Anthropology* (pp. 64–95). Springfield, IL: Charles C Thomas.
- Browne, D.M. (1975). *Principles and Practice in Modern Archaeology*. Bungay: Richard Clay (The Chaucer Press).
- Carver, M. (2009). *Archaeological Investigation*. Oxon: Routledge.
- Carver, M. (2011). *Making Archaeology Happen Design versus Dogma*. Walnut Creek, California: Left Coast Press Inc.
- Carver, M., Gaydarska, B., & S. Monton-Subias. (eds.) (2015). *Field Archaeology from Around the World, Ideas and Approaches*. New York: Springer Briefs in Archaeology. Springer Publishing.
- Congram, D.R. (2008). A Clandestine Burial in Costa Rica: Prospection and Excavation. *Journal of Forensic Science*, 53(4), 793–796.
- Connor, M. (2007). *Forensic Methods Excavation for the Archaeologist and Investigator*. Lanham: AltaMira Press.
- Crist, T.A.J. (2001). Bad to the bone? Historical archaeologist in the practice of forensic science. *Historical Archaeology*, 35(1), 39–56.
- Dupras, T.L., Schultz, J.J., Wheeler, S.M., & Williams, L.J. (2006). *Forensic Recovery of Human Remains Archaeological Approaches*. Boca Raton: CRC Press.
- Foxhall, I. (2000). The Running Sands of Time: Archaeology and the Short Term. *World Archaeology*, 31, 484–498.
- Godwin, G.M. (2001). *Criminal Psychology and Forensic Technology A Collaborative Approach to Effective Profiling*. Boca Raton: CRC Press.
- Greene, K. (1997). *Archaeology An Introduction*. Third Edition. London: Routledge.
- Groen, M., Márquez-Grant, N., & R. Janaway. (2015). *Forensic Archaeology: A Global Perspective*. Chichester: John Wiley & Sons, Ltd.
- Haglund, W.D., Connor, M., & Scott D.D. (2001). The Archaeology of contemporary mass graves. *Historical Archaeology*, 35(1), 57–69.
- Hanson, I. (2004). *The importance of stratigraphy in forensic investigation*. *Forensic Geoscience: Principles, Techniques and Applications*, 232, 39–47.
- Harris, E. (1979). The Laws of Archaeological Stratigraphy. *World Archaeology*, 11(1), 111–117.

- Harris, E. (1989). *Principles of Archaeological Stratigraphy*. Second Edition. London: Academic Press Limited.
- Harris, E. (2002). *Stratification, soil*. In C.E. Orser, *Encyclopaedia of Historical Archaeology* (pp. 598–599). London: Routledge.
- Harris, E.C. (2006). *Archaeology and the ethics of scientific destruction*. In S. Archer & K. Bartoy, *Between Dirt and Discussion, Methods, Methodology and Interpretation in Historical Archaeology* (pp. 141–150). New York: Springer Publishing.
- Hester, T.R. (1997). *Methods of excavation*. In T.R. Hester., K.L. Feder., & H.J. Shafer, Seventh Edition, *Field Methods in Archaeology* (pp. 69–112). California: Mayfield Publishing.
- Hochrein, M.J. (2002). *An autopsy of the grave: Recognising, collecting and preserving forensic geotaphonomic evidence*. In D.W. Haglund., & M.H. Sorg, *Advances in Forensic Taphonomy Method, Theory and Archaeological Perspectives* (pp. 45–70). Boca Raton: CRC Press.
- Hunter, J.R. (1994). *Forensic Archaeology in Britain*. *Antiquity*, 68, 758–769.
- Hunter, J.R., & Cox, M. (2005). *Forensic Archaeology: Advances in theory and practice*. Oxon: Routledge.
- Janaway, R.C. (1996). *The decay of buried human remains and their associated materials*. In J.R. Hunter., C. Roberts., & A. Martin, *Studies in Crime: An Introduction to Forensic Archaeology* (pp. 40–57). London: B.T Batsford Ltd.
- Janaway, R.C. (2002). *Degradation of Clothing and other Dress Materials Associated with Buried Bodies of Both Archaeological and Forensic Interest*. In D.W. Haglund., & M.H. Sorg, *Advances in Forensic Taphonomy Method, Theory and Archaeological Perspectives* (pp. 379–402). Boca Raton: CRC Press.
- Joukowsky, M. (1980). *A Complete Manual of Field Archaeology*. Englewood Cliffs: Prentice Hall.
- Killam, E.W. (2004). *The Detection of Human Remains*. Second Edition. Springfield, Illinois: Charles C Thomas.
- Komar, D., & Buikstra, J.E. (2008). *Forensic Anthropology: Contemporary Theory and Practice*. New York: Oxford University Press.
- Morse, D., Stoutamire, J., & Duncan, J. (1976a). A Unique Course in Forensic Anthropology. *American Journal of Physical Anthropology*, 45(3), 743–747.
- Morse, D., Crusoe, M.D., & Smith, H.G. (1976b). Forensic Archaeology. *Journal of Forensic Science*, 21(2), 323–332.
- Museum of London Archaeology Service. (1994). *Archaeological Site Manual*. Third Edition. London: Museum of London.
- NAS. (2009). *National Academy of Sciences Report Strengthening Forensic Science in the United States: A Path Forward*. Washington D.C.: The National Academies Press.
- Oxford English Dictionary. (2015). *Grave n.1*. <http://www.oed.com/view/Entry/80989?rskey=My2ldX&result=1>. Accessed 16 January 2016.
- Pickering, R.B., & Bachman, D.C. (1997). *The Use of Forensic Anthropology*. Boca Raton: CRC Press.
- Powell, J.F., Steele, D.G., & Collins, M.B. (1997). *Excavation and Analysis of Human Remains*. In T.R. Hester., K.L. Feder., & H.J. Shafer, Seventh Edition, *Field Methods in Archaeology* (pp. 253–282). California: Mayfield Publishing.
- Praetzelis, A. (1993). *The Limits of Arbitrary Excavation*. In E.C. Harris, M.R., Brown, G.J. Brown, *Practices of Archaeological Stratigraphy* (pp. 68–87). London: Academic Press.
- Ramey-Burns, K. (1996). *Forensic Anthropology Training Manual*. New Jersey: Prentice Hall.
- Ramey-Burns, K. (2006). *Forensic Anthropology Training Manual*. Second Edition. New Jersey: Prentice Hall.
- Spennemann, D.H.R., & Franke, B. (1995). Archaeological techniques for exhumations: a unique source for crime scene investigators. *Forensic Science International*, 74, 5–15.
- Tuller, H., & Đurić, M. (2006). Keeping the pieces together: Comparison of mass grave excavation methodology. *Forensic Science International*, 156, 192–200.
- United Nations. (1991). *Manual on the Effective Prevention and Investigation of Extra-Legal, Arbitrary and Summary Executions*. New York: United Nations.
- Wolf, D.J. (1986). *Forensic anthropology scene investigations*. In K. Reichs, *Forensic Osteology: Advances in the Identification of Human Remains* (pp. 3–23). Springfield, Illinois: Charles C Thomas.
- Wright, R., Hanson, I., & Sterenberg, J. (2005). *The archaeology of mass graves*. In J.R. Hunter., & M. Cox, *Forensic Archaeology: Advances in theory and practice* (pp. 137–159). Oxon: Routledge.

Appendix 1 Arbitrary level excavation (after Hanson 2004)



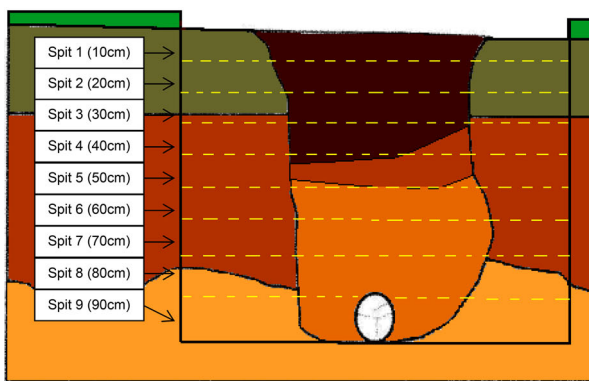
Step 1

Remove O Horizon leaf litter/Turf and Topsoil across a 2 m by 3 m area.

Examine the soil for changes in colouration that can be associated with the mixing of the topsoil and subsoil.

When the graves outline is fully visible, measure, photograph, and map it.

Examine the outline for information about the size and shape of the original digging tools.



Step 2

Remove overburden, including the grave fill in 10 cm arbitrary spits.

Work horizontally, peeling thin layers of dirt and maintain a flat working surface.

If you are suspicious or curious because of changes in density, colour, or texture of the soil, change from a shovel to a trowel to obtain finer control and a sharper, cleaner soil surface.

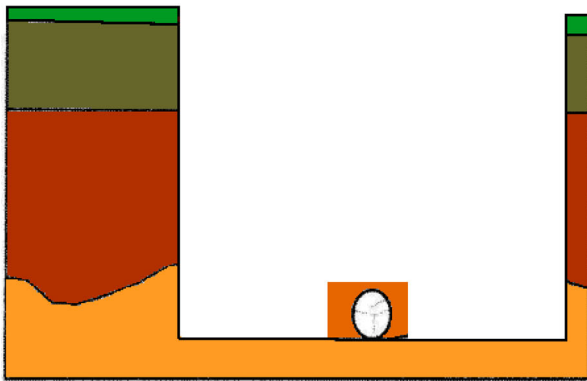
Stop excavating as each feature/object is first discovered. Do not remove it unless it is blocking the progress of the excavation. Even then, only remove the feature after it is exposed in place and completely documented.

If an object appears, change to a brush to avoid tool marks. Objects should be left in place. This leaves them in context with the remainder of the material in the grave for the photographs and maps. As excavation proceeds below the object, it should be left on a soil pedestal (approximately 30 cm in depth), exactly the same size as the artefact, it may expand slightly outward. Do not remove the pedestal unless it is blocking the progress of the excavation.

Sift the soil, spit by spit in sequence. It is not unusual for evidence to be found in the grave fill dirt.

Pedestalling an object

The walls of the pedestal need to be straight under the piece or expand slightly outward. Loose, sandy soils will not hold much weight and the pedestal in such soil may need to be larger than in other soils. Clay soils can hold a lot of water and will shrink as they dry out, so the pedestal may need to be made slightly bigger than the object.



Step 3

Circumscribe the body by digging on all sides to the lowest level of the body (approximately 30 cm). This is similar to digging a ditch around the body. The result takes the form of a pedestal. As stated above, pedestal artefacts in the same way.

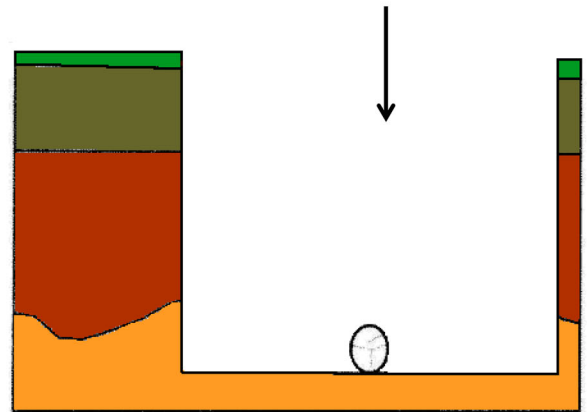
If there is no room to dig around the body, do the best you can. It may be necessary to extend, or (sacrifice) one wall of the excavation to make extra room to manoeuvre on the excavation floor.

Expose the remains and associated evidence by moving laterally, using a soft brush and small tools.

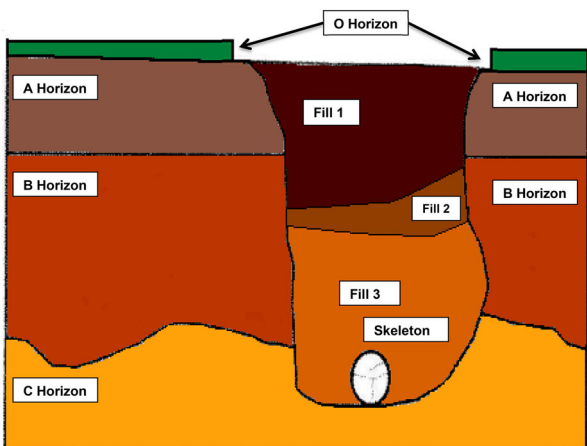
Examine the soil around the skull for hair.

Remove and record all evidence associated with the remains. Some of the evidence may help identify the victim or the perpetrator(s).

Some evidence may aid in the understanding of events around the time of death.



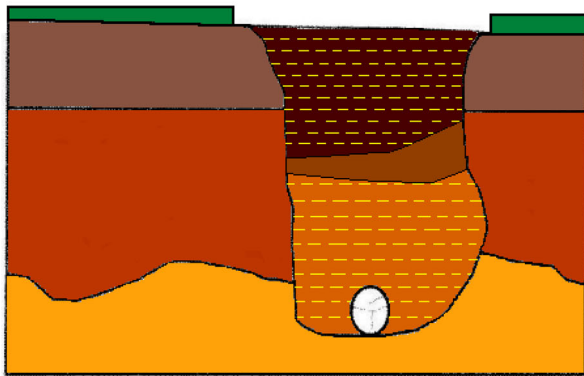
Appendix 2 Stratigraphic excavation (after Hanson 2004)



Step 1

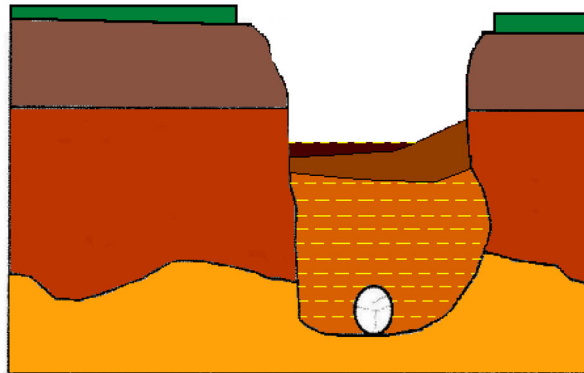
Remove O Horizon leaf litter/Turf and Topsoil to expose the grave in plan.

Note: If turf overlying the grave is loose only remove the loose turf.



Step 2

Carefully remove the grave fill. Ensure that you maintain identifiable stratigraphic boundaries; grave cut(s), different fills etc.



Step 3

Complete removal of grave fill, exposing the skeleton/body and grave surface for analysis.

