Chapter 5: Conclusions

It is evident from the results of this research that a variety of different excavation methods and recording systems are used in the United Kingdom, Ireland, Australasia, and North America to conduct archaeological investigations. To facilitate the use of these different approaches, archaeological organisations operating in these areas have developed archaeological manuals/guidelines through which they have attempted to standardise and improve how their employees conduct archaeological investigations, in order to ensure, that they maximise the accuracy of the recovery and recording processes and produce consistent archives from which they are able to create publications. Although these organisations share the same overall goals, the fact that 153 different archaeological manuals/guidelines are used by different archaeological organisations indicates that there is no singular standardised approach to conducting archaeological investigations. Consequently, this has resulted in multiple excavation methods and recording systems being advocated by different archaeological organisations, and, in turn, led to each of these different methodological approaches to be regarded as standardised approaches.

Despite different archaeological organisations advocating different "standardised" archaeological excavation and recording approaches, there are marked similarities in terms of the methodological approaches that are being advocated between archaeological organisations operating in the United Kingdom and Ireland, and archaeological organisations operating in Australasia and North America. In the United Kingdom and Ireland archaeological organisations define, record and excavate archaeological stratigraphy using similar, if not the same, approaches. Whereas, Australasian and North American archaeological organisations use different excavation methods and recording systems, and define, record, and excavate archaeological stratigraphy differently from British and Irish archaeological organisations, although, use similar approaches to each other. Therefore, there are similarities in the "standardised" approaches being advocated but not on an international scale.

These international methodological differences have been caused as a result of the different archaeological site types that these archaeological organisations have been operating in. Both British and Irish archaeological organisations conduct archaeological investigations on urban sites with complex archaeological stratigraphy, and rural sites

with widely dispersed archaeological evidence with very few complex stratigraphic units present. As a result, the methods that they have developed and advocate in their archaeological manuals/guidelines have been designed to be used and/or adapted for use on both of these archaeological site types. Whereas, Australasian and North American archaeological organisations tend to conduct archaeological investigations on rural Native American or Aboriginal archaeological sites. Such sites contain widely dispersed archaeological evidence and deep stratigraphic deposits, and as a result, the archaeological organisations working on these sites have developed different methodological approaches to investigate such sites, and subsequently, published and advocated these approaches in their archaeological manuals/guidelines.

However, the problem with having multiple "standardised" archaeological excavation and recording approaches advocated within academic literature and archaeological manuals/guidelines, is that archaeologists can apply any one of these techniques to excavate and record the same archaeological feature and justify this by stating that this is a "standardised" archaeological approach. One such example of this was identified during the research's grave excavation experiment, whereby each archaeologist was given the same grave feature to excavate and record, and yet different groups of archaeologists chose to use different excavation methods and recording systems to investigate the grave, each justifying their chosen approach by stating that the approach that they had chosen was standard practice and used regularly during archaeological investigations. However, as little research had been conducted, prior to this research, to comparatively test these different "standardised" excavation methods and recording systems to determine the impact that these different methodological approaches have on data recovery and interpretation-based narratives, advocates of these different approaches had no data from which to state that the methodological approaches that they currently advocate do in fact result in greater accuracy in terms of the archaeological recovery and recording processes as their archaeological manuals/guidelines claim that they do, leaving the question of whether these different "standardised" methodological approaches are equally applicable debatable.

The problem of multiple "standardised" excavation methods and recording systems co-existing and being used during archaeological investigations to excavate the same type of archaeological feature is also present in forensic archaeology. This is due to the fact

that forensic archaeological practitioners have adopted and adapted their methodological approaches from traditionalist archaeological practice. Therefore, forensic archaeological practitioners operating in the United Kingdom and Ireland will use different methodological approaches to their Australasian and North American counterparts, and again, justify their use of these different methodological approaches by stating that they are standard practice in the geographical area in which they are working.

The issue, however, is exacerbated when it is considered in a forensic context, as the primary aim of forensic archaeological investigations is the provision of evidence to legal proceedings. Consequently, when archaeologists are conducting forensic archaeological casework the methods utilised and the evidence retrieved as a result of the investigation are held accountable to the admissibility regulations and the legislative acts of the courts in the country in which the investigation is being conducted. Such legislative acts and admissibility regulations require that any techniques used during the course of a forensic investigation are required to have been – subjected to empirical testing, peer review, have known error rates, have standards controlling their operation, and be widely accepted amongst the academic community from which the methodology originates (Pepper 2005; Glancy and Bradford 2007; Hanzlick 2007; Edwards 2009; Klinker 2009; The Law Commission 2009, 2011; Robertson 2009; Edmond 2010; Robertson 2010; Selby 2010). Therefore, for the evidence that was recovered during the course of the forensic archaeological investigation to be accepted by the court, the forensic archaeologist must be able to demonstrate that the methodological approach that they used during the course of the investigation adhered to a widely accepted and tested forensic archaeological investigatory process (Hunter and Knupfer 1996: 37). However, as previously discussed, no such investigatory process has been established and no substantial empirical testing had been undertaken regarding archaeological excavation methods and recording systems prior to this research, therefore, despite multiple methods being attributed the status of a "standardised" technique, within stringent legal contexts, much of the casework relating to clandestine burials that was undertaken by forensic archaeologists, prior to the completion of this research, failed to meet the admissibility regulations and the legislative requirements of the international court systems.

In order to address these issues this research was undertaken with the aim of determining, which, if any of the various excavation methods and recording systems currently used in the United Kingdom, Ireland, Australasia and North America fulfil the criteria for legal acceptance and best meet the needs of forensic archaeology, focusing on the case of single and mass graves as they are the most common situation faced by forensic archaeologists.

The results of this research indicated that the four archaeological approaches that are currently used to excavate and record cut features, such as graves, are the Stratigraphic Excavation method and Single Context Recording system, the Demirant Excavation method and Standard Context Recording system, the Quadrant Excavation method and Standard Context Recording system, and the Arbitrary Excavation method and Unit Level Recording system. Each of these techniques were subsequently tested against one another using a grave simulation to determine - which of these excavation methods was the most productive and consistent in terms of evidence recovery (evidence included both material evidence and stratigraphic contexts), which of these recording systems provided the most consistent and informative record of the evidence and deposition sequence present in the grave, and which of these excavation methods and recording systems provided the most consistent interpretation-based narrative of the simulated grave's formation process.

The results gained from the grave simulation experiments revealed that the Demirant Excavation method was the most productive in terms of overall evidence recovery retrieving an average of 78.75% of all of the evidence present (Graph 8.8). However, both the Stratigraphic and Quadrant Excavation methods achieved similar evidence recovery rates, retrieving an average of 75.83% of the evidence present (Graph 8.8). In contrast, the Arbitrary Excavation method had an average evidence recovery rate of 39.58% (Graph 8.8). Due to the fact that the average evidence recovery rates were found to be very similar between the Demirant, Quadrant and Stratigraphic Excavation methods and that these differences were proven not to be statistically significant, each of these excavation methods were compared against each other in terms of the consistency of evidence recovery. The data obtained indicated that the Quadrant Excavation method was the most consistent with a variation rate of 33.33%, followed by the Demirant Excavation method with a variation rate of 41.67%, and finally, the

Stratigraphic Excavation method with a variation rate of 50% (Graph 8.8). These results suggest that the Quadrant Excavation method is the most reliable excavation technique out of the four methodological approaches tested, and therefore, the most suitable excavation method to use in forensic archaeological casework.

In terms of which recording system produced the most consistent and informative record of the evidence and deposition sequence present within the simulated grave, each recording system provided a systematic approach to document both the evidence and deposition sequence present. However, the Standard Context Recording system associated with the Quadrant Excavation method was deemed to be the most informative and consistent recording system used. This was due to the fact that through using this recording system the participants produced both a long section and half section drawing of the simulated grave. This meant that the records that they produced illustrated each of the context's dimensions in their entirety across the length and width of the simulated grave, in a clear and understandable manner, which is advantageous in forensic contexts as it means that lay jurors would be able to easily understand the deposition sequence of the grave. Furthermore, as such recording can be done in the field, it requires less time in terms of editing after the investigation has been completed, saving the investigation both time and money, which are crucial variables to be taken into consideration for work that is conducted within a forensic context.

The Standard Context Recording system associated with the Demirant Excavation method also proved to be systematic and thorough. However, the problem with this recording system is that it relied on using a half section drawing to illustrate the deposition sequence of the simulated grave, and as some of the contexts sloped and failed to reach the half section point, these contexts did not appear in the half section drawing. Therefore, participants using this recording system were required to draw additional plans of the contexts that did not reach the half section point. Additionally, the half section drawing only displayed the deposition sequence of the simulated grave at one point, and therefore, failed to illustrate the dimensions of individual contexts across the entire simulated grave. This meant that the participants using this recording system were required to spend additional time, after the investigation had been completed, reconstructing the simulated grave's deposition sequence using the

measurements of the contexts that were noted down on their context sheets, thus costing the investigation both more time and money.

The Single Context Recording system associated with the Stratigraphic Excavation method also produced a comprehensive set of records relating to the evidence and deposition sequence present within the simulated grave. However, the planning process associated with this technique was particularly time consuming, and required further post-investigation processing to superimpose each of the separate context plans that were produced in order to illustrate the deposition sequence of the simulated grave in its entirety. Consequently, costing the investigation both more time and money.

The Unit Level Recording system associated with the Arbitrary Excavation method proved to be an effective recording system to use to document the location and relative deposition sequence of material evidence. However, the Arbitrary Excavation method associated with this recording systems use resulted in participants digging through and destroying the dimensions of individual contexts, particularly those with any gradient in their deposition. Therefore, this recording system resulted in the production of a list of artefacts and their spatial location without any accurate contextual information. Consequently, this recording system provided an inaccurate reconstruction of the grave simulation's deposition sequence and must be deemed as uninformative and unsuitable for use in forensic archaeological casework.

In regards to which excavation method and recording system provided the most consistent interpretation-based narrative of the simulated grave's formation process, it was evident from the results of this research that the accuracy of the narratives that were produced was directly correlated with the excavation method and recording system that the participants chose to use. Therefore, those participants that chose to use the Stratigraphic, Demirant and Quadrant Excavation methods to excavate the grave simulation recovered the greatest amount of evidence, and subsequently, their narratives were the most accurate. Whereas, the participants who used the Arbitrary Excavation method recovered the least amount of evidence, and therefore, their narratives were the least accurate. Consequently, as the Quadrant Excavation method produced the most consistent results in terms of both the evidence that was recovered and the records that were produced, the participants using this approach also produced the most consistent interpretation-based narratives of the simulated grave's formation sequence.

An additional variable taken into consideration when comparatively evaluating each of the different excavation methods and recording systems was the length of time that it took participants to complete the grave simulation excavation experiment using each of the different approaches. The results indicated that on average, the Stratigraphic Excavation method took 5 ½ hours to complete, the Demirant Excavation method took 5 hours to complete, the Quadrant Excavation method took 4 ½ hours to complete, and the Arbitrary Excavation method took 2 ½ hours to complete (Table 8.166; Graph 8.10). The reason why greater lengths of time were spent excavating and recording using the Stratigraphic, Demirant and Quadrant Excavation techniques was because these methodological approaches required participants to spend time delineating the boundaries of the individual contexts contained within the grave simulation, and also required more in depth recording, including - the completion of context sheets, plans and section drawings. In comparison, the Arbitrary Excavation method required no such skill as the contexts contained within the simulated grave were removed in set 10cm spits with no attention paid to the presence of multiple contexts or such contexts' dimensions. Overall, due to the fact that the length of time that was spent excavating and recording using the Stratigraphic, Demirant and Quadrant Excavation methods varied very little, it suggests that the utilisation of any one of these methods during the course of a forensic archaeological investigation will not require a significant increase in the amount of time needed to complete the investigation. Therefore, when forensic archaeologists are deciding on which methodological approach to use, time should not be considered as an exclusionary factor during the decision making process.

The length of time that the participants spent excavating and recording the grave simulation was also analysed to determine if time had any impact upon the amount of evidence that was recovered. The results indicated that there was a minimal correlation between the length of time that the participants spent excavating and recording and the amount of evidence that they recovered. Therefore, a greater length of time spent excavating and recording a grave will not necessarily result in an improvement in the quality and quantity of evidence that is recovered during a forensic archaeological investigation. Rather, other factors, such as the methodological approaches that are used and the observation skills of the archaeologist will determine whether evidence will be successfully recovered during a forensic archaeological investigation.

An additional objective of this research was to evaluate the impact that archaeological experience had on the quality and quantity of evidence recovered during the grave simulation excavation experiment. The results indicated that there was no correlation between archaeological experience and evidence recovery rates, and that archaeological experience can only be used to predict an archaeologist's ability to achieve a higher rate of evidence recovery in 4% of cases (Graph 8.14). The reason why archaeological experience had little impact on evidence recovery is due to the fact that experienced archaeologists have obtained significant experience in excavating a variety of different archaeological features. They then use this experience to predict what an archaeological feature, such as a grave, might contain and how it may be constructed, and presume that the archaeological feature that they are excavating will match similar archaeological features that they have excavated in the past, resulting in a complacent attitude and approach to the investigatory process. Inexperienced archaeologists, however, are less confident in their abilities and are resultantly more cautious and considerate when conducting archaeological investigations, to ensure that they have not missed any evidence or misinterpreted what they have uncovered.

As a result of this research, each of the four different methodological approaches tested during the simulated grave excavation experiments now satisfy the legislative and admissibility requirements of the international courts. Although, on the basis of the findings of this research, it is apparent that the Quadrant Excavation method and Standard Context Recording system best meets the needs of forensic archaeology. However, the Demirant Excavation method and Standard Context Recording system was also highly productive in terms of evidence recovery, though due to the weaknesses of the recording system, it is suggested that this approach should only be used in situations in which the Quadrant Excavation method is unable to be applied, for example, when excavating particularly small clandestine burials or when excavating in loose soils, as the quadrants associated with the Quadrant Excavation method have a tendency to collapse under such conditions. Moreover, the Stratigraphic Excavation method and Single Context Recording system was also reasonably successful in regards to evidence recovery, but due to the fact that this approach produced inconsistent results and that the recording procedure was highly time consuming, it should only be used in situations in which the two aforementioned techniques are unable to be applied. The one

approach that should not be used in forensic archaeological casework is the Arbitrary Excavation method and Unit Level Recording system. This is because this methodological approach had an extremely poor evidence recovery rate, and destroyed the deposition sequence present within the simulated grave. Moreover, through statistical testing, this methodological approach was found to recover significantly less evidence than the three other archaeological techniques tested and the Control participants. Consequently, as this methodological approach is highly unreliable and produces inaccurate data, any forensic archaeological reports that have been produced from forensic archaeological investigations that have used this approach, should be treated with extreme caution.

Given the extent to which the Arbitrary Excavation method and Unit Level Recording system has been shown to be used during archaeological fieldwork outside the United Kingdom and Ireland, the findings of this Project are particularly alarming. It suggests that archaeological data, whether it has been obtained during a research, commercial or forensic archaeological excavation, using this methodological approach, is incomplete, and any archaeologists that have relied on this data to formulate interpretations about a site are likely to have produced inaccurate or at least questionable interpretations, as they have used a data collection which has, on average, 60.42% of its data missing (Graph 8.8). In light of this finding, scholars must be cautious of any archaeological reports or interpretations that have relied on data collections obtained using the Arbitrary Excavation method and Unit Level Recording system, and must focus their attention on re-assessing this archaeological data in light of the findings of this Project.

Another cause for concern amongst the archaeological community is that this Project demonstrated that archaeological data recovery rates varied greatly, between archaeologists in general with archaeological experience proving to have little bearing on data recovery rates, and within individual methodological approaches. This lack of consistency both within methods and between practitioners suggests that one cannot assume that an experienced team of archaeological practitioners will necessarily retrieve the maximum amount of archaeological data available, which is the ultimate aim of any archaeological investigation. It is therefore necessary, for a skills test to be established that will ensure that archaeological practitioners are consistently producing high quality archaeological investigations, and that the interpretations that are produced on the basis

of these excavations are reliable. It is particularly important that skills tests are introduced into the sub-discipline of forensic archaeology, as the consequences of forensic investigations can be profound.

Chapter 6: Recommendations

Specific recommendations for forensic archaeological investigations:

In light of the findings of this research, when conducting forensic archaeological investigations the Quadrant Excavation method and Standard Context Recording system should be used. If this approach is unable to be utilised the Demirant Excavation method and Standard Context Recording system, or the Stratigraphic Excavation method and the Single Context Recording system should be used. Any deviation from these recommended approaches should be justified in the forensic archaeologist's report.

When using the Quadrant Excavation method or the Demirant Excavation method all recording forms should be completed as the excavation progresses, not only after a half or quadrant has been removed. This will prevent any material evidence from later being incorrectly associated with the wrong context.

When conducting forensic archaeological investigations, all spoil related to different contexts should be stored separately. All spoil should also be sieved, either at site or in controlled laboratory conditions. This will ensure that any material evidence items that were missed during the excavation process are recovered and can be reassociated with the context from which they originated.

General recommendations:

An internationally recognised forensic archaeology organisation/committee should be formed in order to share ideas, research, case studies, and methodological developments within the sub-field. The formation of such an organisation/committee will improve the standard and consistency of forensic archaeological practice on an international scale. Furthermore, when forensic archaeologists of different nationalities are deployed to work on international projects, such as mass grave excavations, they will then all be familiar with the internationally advocated methodological approaches, and in turn, this will help to prevent disagreements between forensic archaeologists about which methodological approaches should be used.

An internationally recognised forensic archaeology investigation protocol should be developed for single and mass graves. The development of such a protocol could be informed by the results of this research, and could be facilitated by the formation of an internationally recognised forensic archaeology organisation/committee.

A forensic archaeology competency test should be developed for archaeologists who wish to work in forensic contexts. Such a test would require archaeologists to excavate and record a controlled grave simulation, such as the one used in this research. In addition, as archaeological experience proved to have little bearing on evidence recovery rates, such competency tests should be repeated every five years in order to ensure that the highest archaeological standards are being maintained. The implementation of such competency tests would be facilitated by the formation of an internationally recognised forensic archaeology organisation/committee.

The applicability of the Quadrant, Demirant and Stratigraphic Excavation methods and their associated recording systems should be tested on mass graves. This will determine whether these methodological approaches are as suited to mass grave investigations as they are to single inhumation investigations.

All archaeological sites that have been investigated using the Arbitrary Excavation method and Unit Level Recording system need to be re-examined and interpretations formed using data collections obtained through the use of this method need to be re-evaluated.