An epistemology of archaeological excavation. A comparative study of British field practice.

Gustavo Sandoval García

PhD

University of York

Archaeology

September 2019

Abstract

In recent decades, there has been a trend for increasing the number of written formats in excavation records. For example, this has implied the revival of excavation diaries and calls for the use methods based multiple types of standardized sheets (context and feature sheets). These ideas are grounded in the methodological premise that more complex recording systems will enable to produce higher-quality records that will include more interpretative and reflexive data. In general, these ideas have been widely considered in Britain where there is an important debate about field-methods. However, there is an important flaw in these views for being inconsistent with interpretative theory which suggests that interpretation and reflexivity depends on the epistemic context of an investigation.

This study examined the primary records of the three British projects that implemented alternative recording strategies. These include, the excavation diaries from Catalhöyük recorded under the principles of the reflexive method. The context sheets from a commercial site in London. And the context and feature sheets from the explorations of a medieval site in Scotland, following the principles of the feature system. The goal of my investigation was to compare these sources to investigate whether the quality of records improved with a more complex method. Results show that reflexive diaries contain additional information not observed in other systems, but a lot of it is trivial. On the contrary, the analysis of context and feature sheets shows that these formats include multiple types of interpretative and reflexive information, but frequently these data are not reported within an adequate standard. Moreover, documental analysis demonstrates how the epistemic context of projects affects their interpretative results.

Overall, this study creates a more promising scenario for an interpretative science that usually puts more confidence in the standardization and instrumentalization of practices rather than in the expertise of fieldworkers. Finally, my investigation suggests that there are good elements to re-formulate interpretative theory not only as an individual-based practice, but also as a collective exercise in which the interaction among team members is crucial for the investigation of sites.

Contents

Abstract ... 3
List of tables ... 10
List of figures ... 11
List of graphics ... 16
List of accompanying materials ... 20
Acknowledgments ... 21
Author's declaration ... 22
Abbreviations ... 23
Introduction ... 24
Reflexivity and the standardization of recording methods ... 25
The feature system and the instrumentalist notion of recording methods ... 28
The problem of quality in excavation records. ... 30
The importance of the epistemic context in excavation work ... 31
Description of the thesis ... 33
Chapter 1 Interpretation in archaeological excavation ... 35
1.1 A conceptual distinction ... 37
1.2 Observation and skill ... 41
1.3 Techniques and standardization ... 44
1.4 The act of discovery and the expertise of fieldworkers ... 49
1.5 Background knowledge and experience ... 55
1.6 The character of remains ... 57
1.7 The role of excavation strategy ... 60
1.8 Collective empiricism: The team and the record ... 61

Chapter 2 Methodology.....67
2.1 Case selection.....68
2.2 Sources.....70
2.3 Internal and comparative analysis of case studies.....72
Coding.....72
Quantitative observations: variability and frequency.....72
Qualitative observations: relevance and validity of data.....73
Comparative analysis.....74
2.4 Basics of coding.....74
A conceptual background for interpreting the meaning of records.....76
A contextual strategy for interpreting the meaning of records.....77
Atypical meaning problems.....77
2.5 A classificatory framework for sources.....79
Non-interpretative information.....81
Low level claims: Contexts descriptions and finds reports.....82
Stratigraphic claims: layer contacts, stratigraphy and correlations.....84
Assemblage: group of finds.....85
Structural elements.....86
Features: group of contexts.....87
Group of features, Spaces, Buildings and Middle-level contrasts.....88
Formation processes.....90
Phases.....91
High level claims: Function, Agency and Culture/Period.....91
Chapter 3 Reflexive archaeology.....96

3.1 The reflexive method 97

3.2 The project and the site 100

3.3 The sample 104

3.4 The narrative format of reflexive diaries 105

3.5 Results....106

3.5.1 Non-Interpretative information 109

3.5.2 Contexts 111

3.5.3 Finds....113

3.5.4 Other Low-level claims....115

3.5.5 Stratigraphic information 115

3.5.6 Assemblages features and structural elements 117

3.5.7 Formation processes 120

3.5.8 Building and spaces 122

3.5.9 Constructive phases 124

3.5.10 High-level claims....125

3.6 Reconsidering reflexivity....126

3.7 The performance of fieldworkers....132

Empowerment, authority and responsibility of reflexive diggers 135

3.8 Longitudinal analysis....136

3.9 General observations 138

Chapter 4 Single context recording....141

4.1 The recording system: a sequence of related events 142

4.2 The analysis of contexts sheets 144

4.3 The project and the site: MRG95 150

4.4 The sample 152

4.5 Results..153

4.5.1 Non-interpretative information..156

4.5.2 Contexts 159

4.5.3 Finds and samples..163

4.5.4 Other Low-level claims..164

4.5.5 Stratigraphic information 165

4.5.6 Correlations 168

4.5.7 Structural elements 170

4.5.8 Features 175

4.5.9 Formation processes 179

4.5.10 Groups of features, spaces, buildings and phases..182

4.5.11 High-level claims...187

4.6 The importance of documental analysis. 190

4.7 The performance of fieldworkers..193

4.8 Longitudinal analysis..198

4.9 General observations 201

Chapter 5 . The feature system 204

5.1 The recording system..205

5.2 The project: The Tarbat Discovery Programme..210

5.3 The site..212

5.4 The sample 214

5.5 Results...215

5.5.1 Non-Interpretative information 217

5.5.2 Context descriptions 219

5.5.3 Finds and samples 224

5.5.4 Other low-level claims 225

5.5.5 Stratigraphic information 225

5.5.6 Correlations 227

5.5.7 Assemblages 228

5.5.8 Structural elements 229

5.5.9 Features 234

5.5.10 Formation processes 239

5.5.11 Groups of features, Structures, Spaces and Buildings 246

5.5.12 Phases 251

5.5.13 High Level claims 252

5.6 Digger’s performance 255

5.7 Longitudinal analysis 259

5.8 General observations 260

Chapter 6 . Discussion 263

6.1 Comparation by type and epistemic context 264

6.2 Reflexivity in excavation records 267

6.3 Decisions and observational circumstances in excavation records 270

6.4 Credibility-of testimonies in excavation sheets 272

6.5 A final contrast. Virtues and problems in recording methods. 278

6.6 Visual and textual testimony in site records. 281

6.7 Individual and collective interpretation in the field. 285

6.8 Future directions 289

Concluding remarks 293

References...298

List of tables

Table 2.1. The table resumes the main criteria for case selection. 70

Table 2.2 Classificatory framework for excavation records. 93

Table 4.1. Participation in the project by digger. MRG95.. 194

Table 5.1. Participation in the project of the ‘main team’. Tarbat.. 256

List of figures

Figure 1.1. A site of archaeological dissection. Burdale, Yorkshire. (Copyright Julian Richards CC BY 4.0). https://doi.org/10.5284/1021540 45

Figure 1.2. A site of archaeological dissection. North area excavation. Catalhöyük, Turkey. (Copyright Catalhöyük Research Project. CC BY-NC-SA 2.0). 45

Figure 1.3. The discipline of recording primary testimonies onsite. Burdale, Yorkshire. (Copyright Julian Richards CC BY 4.0). https://doi.org/10.5284/1021540 48

Figure 1.4. Exposed buildings in the North Area, Catalhöyük. (Copyright. Catalhöyük Research Project. CC BY-NC-SA 2.0). 53

Figure 1.5. Two excavators removing a fill to define a cut. Burdale, Yorkshire. (Copyright. Julian Richards CC BY 4.0). https://doi.org/10.5284/1021540 53

Figure 1.6. Mesolithic structure at Star Carr. Site photography (left), plan drawing (right). (Copyright Star Carr Project CC BY 4.0). https://doi.org/10.22599/book1 59

Figure 1.7. A group of diggers defining a feature. Heslington East, University of York (Copyright York Archaeological Trust, CC BY 4.0). https://doi.org/10.5284/1019860 64

Figure 1.8. Priority tours for discussion site interpretations at Building 49, Catalhöyük. (Copyright. Catalhöyük Research Project. CC BY-NC-SA 2.0) 64

Figure 2.1. An illegible testimony of a context sheet 78

Figure 2.2. Sketch from a context sheet (MW-710). 79

Figure 2.3. The recovery of finds from a fill and the definition of a cut. Burdale, Yorkshire. (Copyright Julian Richards CC BY 4.0). https://doi.org/10.5284/1021540. 83

Figure 2.4. Representation of stratigraphic relations (A) and correlations (B) (above) and layer contacts (below). Source: Harris, 1989: 36. 84

Figure 2.5. Defining a brick floor. (Copyright Oxford Archaeology CC BY 4.0). https://doi.org/10.5284/1047571 86

Figure 2.6. Types of features. Source: Carver, 2009: 21. 87

Figure 2.7. Defining spaces of Building 77, Catalhöyük. (Copyright. Catalhöyük Research Project. CC BY-NC-SA 2.0). 88

Figure 2.8. Types of activity areas. Source. Flannery, 1976: 45. 89

Figure 3.1 Excavations at East and West mound of the Catalhöyük Research Project (Copyright Catalhöyük Research Project)....101

Figure 3.2 The North area (Copyright Catalhöyük Research Project)....102

Figure 3.3. The final constructive phase of B49 (Copyright Catalhöyük Research Project CC BY-NC 4.0). 103

Figure 3.4. Detection of B49 (Copyright Catalhöyük Research Project CC BY-NC 4.0)....104

Figure 3.5. Building 49 (Copyright Catalhöyük Research Project)....123

Figure 4.1. Frontside of a context sheet describing the interpretive process of a fill of a wooden lined pit (SC-123). The image also illustrates the coding process of a context sheet into different types of claims. 147

Figure 4.2. The front and back side of a context sheet describing a timber lining (SC-124)....148

Figure 4.3. The front side of a context sheet describing a cesspit cut (HK-134). 149

Figure 4.4. Site location in relation to Roman London map (Copyright MOLA)....150

Figure 4.5. MRG95. Furnace complex. Roman period, Phase 1 (Copyright MOLA)....151

Figure 4.6. Nuclear area of furnace complex. Roman period, Phase 1 (Copyright MOLA). 151

Figure 4.7. Observational conditions at MRG95 (Copyright MOLA)....157

Figure 4.8. Sketch from a partially excavated cut (MW-1056)....157

Figure 4.9. Sketches describing operative decisions. JA-1759 (left) and HLB-1047 (right)....159

Figure 4.10. Sketch of a layer (SC-1424)....161

Figure 4.11. Sketches depicting ambiguous layer boundaries. MLC-856 (left) and MLC-1030 (right)....162

Figure 4.12. Sketches of timber stakes. GT-1227 (left), GT-1249 (centre), GT-1259 (right). ...164

Figure 4.13. Sketches depicting cutting relation. SC-974 (left), ME-1119 (right)....166

Figure 4.14. Examples of uncertainty claims in the stratigraphic sequence. 167

Figure 4.15. Corrections in the sequence over related sheets..167

Figure 4.16. Corrections over the sequence. 168

Figure 4.17. Sketches depicting correlated contexts. MLC-321 (left), SC-1115 (right). 169

Figure 4.18. A preliminary sketch for layer [1010] (left) and the updated version (right). (PGF/MLC-1010)..170

Figure 4.19. The vaulted floor from the red kiln (S3). (Copyright MOLA). 171

Figure 4.20. Sketch of a posthole. LAB-1656 (left) and sketch of partially discovered kiln. MW-871 (right)...173

Figure 4.21. A series of related sketches depicting various contexts of a posthole. An excavated cut with the mark of a stone slab (left). The upper surface of a fill of stones overlying the cut (centre) and a squared slab in the base of the cut (right). JA-1421 (left) and JA-1437 (centre and right)...174

Figure 4.22. Sketches of structural remains. L'ON-1464 (left), SC-1372 (right). 175

Figure 4.23. Sketches of features. LAB-1026 (left) and LAB-1054 (right)..............................175

Figure 4.24. Cut of a structural feature. AH-1000 (left) and sketch of a non-structural feature. MW-1452 (right)..176

Figure 4.25. Sketches depicting uncertain situations of feature definition. HK-758 (left), PGF/MLC-804 (right). 178

Figure 4.26. A sketch representing an ambiguous feature relation between a layer and a cut (MW-1055)...179

Figure 4.27. Sketches of groups of features. AH-988 (left), PFG-991(right). 183

Figure 4.28. Sketches of groups of features. LAB-1316 (left), HK-254 (right).........................184

Figure 4.29. MRG95. Roman Period. Phase 2 (MOLA copyright)..186

Figure 5.1. Context sheet of the recording system used at Tarbat. Courtesy of FAS.................206

Figure 5.2. Feature sheet of the recording system used at Tarbat. Courtesy of FAS. 209

Figure 5.3. Excavation sectors at Tarbat (Copyright Tarbat Discovery Programme). 211

Figure 5.4. General view of Sector 2 (TRS2) (Copyright Tarbat Discovery Programme). 211

Figure 5.5. General plan of the monastic occupation in Sector 2 (Copyright Tarbat Discovery Programme). 213

Figure 5.6. Sketch of a posthole describing the operative decisions regarding its exploration (F279-LZ). 219

Figure 5.7. The description of a layer as visual estimates. (C1360-KJA/DW). 220

Figure 5.8. Context description for a dump of fired clay (C2547) related to a hearthe (F148). (C2547-NJT). 223

Figure 5.9. Stratigraphic diagrams in feature sheets. F299-JMM (left) and F378-CKH (right). 226

Figure 5.10. The stratigraphic section of a context sheet (C1146-TJS). 226

Figure 5.11. A fragment of F460 showing stratigraphic information. (F460-KM). 227

Figure 5.12. Sketch illustrating two types of correlations. One for F23=F99 excavated in two interventions and another for F52=F90 separated by F2. (F23-RTJ/ST). 228

Figure 5.13. Sketch of a cow burial showing the spatial arrangement of bones (C1734-CAS)..229

Figure 5.14. Sketch of a section for road F469. (C3522-LAC) 231

Figure 5.15. Feature sheet of F148. The feature was initially described by ST in 1999, but the description continued and concluded in 2004 by NJT. 233

Figure 5.16. The sketch illustrates and discusses the unclear boundaries between F52 and F60 (F52-JB/ST). 238

Figure 5.17. Feature sheets of a pit with an unreliable information of disuse events and failing to interpret constructive and use events (F345-LG). 241

Figure 5.18. Section for constructive evidence of a feature sheet. (F52-JB/ST). 242

Figure 5.19. Sketch of the possible depositional process of layer C2447. (F395-LG) (left) and a sketch of the cut related to the constructive process of wall F97 (right). (F97-CKH)................243

Figure 5.20. Feature sheet describing a flue and its depositional history (F395-LG)................245

Figure 5.21. A group of stake holes defining a timber structure around hearth F535 (Copyright Tarbat Discovery Programme). https://doi.org/10.5284/1031216 247

Figure 5.22. Metalworking area defined by the clustering of three hearth features F148, F353 and F479 (Copyright Tarbat Discovery Programme). https://doi.org/10.5284/1031216 247

Figure 5.23. Structure (S9) or the vellum workshop. (Copyright Tarbat Discovery Programme). https://doi.org/10.5284/1031216 248

Figure 5.24. Gateway access defined by F436, F442 and F443. (F443-CKH). 250

Figure 5.25. Preliminary sketch from an excavation diary representing an occupation phase at TRS2 defined. 252

Figure 5.26. Section record for F345 (LG-F345). 255

Figure 6.1. Example of a problematic context sheet that fails to stablish the type of layer being described (LO'N-202). 276

Figure 6.2. Example of an unreliable context sheet that claims the secondary character of the main backfill of F345 but failing to explain the relation between evidence and the formation process (C1827-LG). 277

List of graphics

Graphic 3.1. General distribution by type. B49. The graphic illustrates the variability and percentage of information types in excavation diaries. 107

Graphic 3.2. General distribution and relevancy by class. B49. This graphic illustrates the distribution of information by class/subclass and the percentage of trivial information for non-interpretative and reflexive classes. The proportion of trivial reflexivity has been calculated in relation to the percentage of interpretative reflexivity. This includes all reflexivity types except Methodology (blue area of the column). Middle level A summarizes assemblage, structural elements, features and formation processes claims. Middle level B summarizes building, space and phase claims. High-level summarizes functional, agency and culture/period claims. 108

Graphic 3.3. Reflexivity B49. This graphic illustrates the dstribution of reflexive claims in relation to three qualitative categories (evidential, observational and persistent reflexivity) and interpetive types. The graphic also includes the proportion of problematic testimonies for evidential testimonies. 109

Graphic 3.4. Distribution of the sample by fieldworker, B49. The graphic illustrates the participation of fieldworkers in diary recording and indicates their professional status: professional diggers (P), graduate students (GS) and trainees (t). 133

Graphic 3.5. Diggers' performance B49a. The graphic illustrates the distribution of various types of claims by fieldworker. These types seem to be significative to indicate differences in the interpretative performance of fieldworkers produced by skill and participation in the project. .134

Graphic 3.6. Diggers' performance B49b. The graphic illustrates the distribtion of different types of cliams by fieldworker. This distribution is considered to be consequence of the type of remains excavated by participants. 135

Graphic 3.7. Distribution of the sample by season. B49. 137

Graphic 3.8. Longitudinal analysis. B49. The graphic illustrates the distribution of claims across the seasons of the project. 138

Graphic 4.1. Distribution of the sample. MRG95. The graphic illustrates the number of sheets collected by each hundred of the site's sequence. 152

Graphic 4.2. General distribution by type. MRG95. The graphic illustrates the frequency and percentage by type of claim. Frequency indicates how often a claim was recorded and percentage indicates how much information of a type was coded. 154

Graphic 4.3. Credibility of interpretative claims. MRG95. This graphic describes the distribution of valid claims and invalid claims for interpretaive information. It also includes the proportion of best practice tetsimonies for non-interpretative data (circumpstances and decisions). 155

Graphic 4.4. Reflexivity. MRG95. The graphic illustrates the distribution of reflexive claims into three categories (observational, persistent and evidential) and its relation to interpetive aspects. Also, it illustrates the frequency and distribution of recording problems for evidential claims. 156

Graphic 4.5. Distribution of the sample by digger. MRG95. This graphic illustrates the proportion of context sheets recorded by each participatn in relation to the extent of his participation in the project. 195

Graphic 4.6. Digger's performance. MRG95a. This graphic illustrates the percentage of valid and invalid basic interpretations by fieldworker. Basic interpretations include context, structural and formation porcess data. Additionally, the graphic illustrates the distribution of problematic reflexive testimonies, relevant reflexive claims, inferences and functional claims...................196

Graphic 4.7. Digger's performance. MRG95b. This graphic illustrates the distribution of different types of claims by fielworker. 197

Graphic 4.8. Distribution of the sample into three arbitrary stages. MRG95. 199

Graphic 4.9. Longitudinal distribution. MRG95. This graphic shows the distribution of valid and invalid interpretations, and various types of claims over the three phases of the project. Each phase is defined by a period of five weeks. 200

Graphic 5.1. Distribution of the sample. Tarbat. 214

Graphic 5.2. General distribution by type. Tarbat. The graphic illustrates the frequency of claims in context and feature sheets. And the percentual distribution of data in the sample. 216

Graphic 5.3. Credibility of interpretive claims. Tarbat. The graphic illustrates the distribution of valid and invalid interpretative data for the total sample (General Index) and by type. Additionally, it includes the describes the proportion of best practice claims for circumstantial and decision claims. Finally, it depicts the proportion recording mistakes for some types of claims. 216

Graphic 5.4. Reflexivity. Tarbat. The graphic illustrates the distribution of reflexive claims according to three qualitative categories (observational, evidential and persistent) and in relation to interpretive claims. Also, it describes the distribution of recording problems in evidential claims. 217

Graphic 5.5. Distribution of the sample by fieldworker. The graphic also indicates the distribution of context and feature sheets by participant. Main team {RBM…LG} and Backup team {AR...KJA}. Professional diggers (p), trainees (t) and volunteers (v) 257

Graphic 5.6. Diggers' performance. Tarbat. The graphic illustrates the distribution of valid and invalid basic interpretations for the participants of the main team (RBM…DMK) and some participant of the 'backup team' (LAC-KJA). The graphic also includes the distribution of Inferences, Evidential reflexivity, Space, Group of features and Functional claims.................258

Graphic 5.7. Distribution of the sample into three arbitrary stages. TRS2. 259

Graphic 5.8. Longitudinal distribution. Tarbat. This graphic illustrates the distribution of valid and invalid basic interpretations, and various types of claims over the course of three arbitrary stages of the project..260

Graphic 6.1. Comparation by type and case study. The graphic illustrates the distribution of the most recurrent types of claim in projects. Reflexivity summarizes the percentage of uncertainty, fluidity and possibility claims...265

Graphic 6.2. Comparation of reflexive claims by case study. The graphic illustrates the percentage of observational, evidential and persitent claims. Observational claims only include intentional testimonies like narratives of discovery and sketches to represent uncertainty and fluidity. Unintentional testimonies like corrections over stratirgaphic diagrams are not counted. 270

Graphic 6.3. Comparation of non-interpretive claims by case study. Systematic non-interpretative and Best practice non-interpretive categories summarize circumstantial and decision claims by project. 271

Graphic 6.4. Comparation of the credibility of testimonies by case study. This graphic illustrates the general proportion of valid and invalid testimonies in the total sample of projects. This evaluation is based in the reliability of basic claims (layer description, structural and formation process claims). The graphic also includes the frequency of invalid layer descriptions and formation process claims. 274

List of accompanying materials

The accompanying CD includes three files.

Three files with digital copies of the coding process of records. One for each case study.; these files are formatted to run with Nvivo (.nvp).

- Case 1. B49.
- Case 2. MRG95
- Case 3. Tarbat

Acknowledgments

This investigation was funded by the people of Mexico through the Consejo Nacional de Ciencia y Tecnología (Conacyt). To them, my most sincere gratitude.

Additionally, this research was supported by numerous people and institutions. Karen Thomas (MOLA) and Veronica Fraser (HES) facilitated the access to archaeological archives at London and Edinburgh. Equally important was the advice of people who generously shared information about projects and recording systems: James S. Taylor, Daniel Eddisford, Colleen Morgan, Heather Knight, Cecily Spall and Kurt-Hunter Mann. Kevin Walsh, Martin Carver and Steve Roskams contributed more systematically in the development of this investigation. Finally, I wish to thank Chris Evans and Penny Spikins for revising the manuscript. Overall, if there is something good in this investigation is for being nourished by the rich tradition of British fieldwork.

I would like to thank my family and friends for their support, but they deserve something better than scholarly claims.

Author's declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Abbreviations

Recording formats and systems

CS	Context sheet(s)
FS	Feature sheet(s)
SCR	Single context recording

Projects

CRP	Catalhöyük Research Project
NGH	Northgate house
TDP	Tarbat Discovery Programme

Sites

B49	Building 49, Catalhöyük
MRG95	Moorgate street 20-28, London
Tarbat	Tarbat, Sector 2

Institutions

FAS	Field Archaeology Services
HES	Historic Environment of Scotland
LAARC	London Archaeological Archive
MOLA	Museum of London Archaeology

Introduction

Since the establishment of PPG16 in 1990, the archaeology of Britain has been practiced in two sectors broadly defined as the academic and the commercial. In one hand, the academic sector includes the archaeological research produced by scholar bodies as universities and institutes. On the other, the commercial sector includes archaeological research made under development-led projects, formerly known as rescue archaeology. Commercial archaeology follows a free-market model based in competitive tendering with private companies as providers of archaeological services (Carver 2009, 366). This is an alternative model to other forms of development-led archaeology in which the government regulates and provides archaeological services. For instance, the INRAP is a national institute that centralizes preventive archaeology in France (Kristiansen 2009, 642; Demoule 2002, 175).

A recurrent critic towards the British model is that free market has had negative effects in the quality of research because archaeological work is mainly assigned in relation to price (Demoule 2002, 173; Schlanger and Aitchison 2010; Demoule 2011). In Britain, commercial archaeology has been received ambivalently. Some people describe it as an improvement of practice due the professionalization of diggers and standardization of methods (Jennings 2011, 283). Besides, the quantity of projects and the amount of funding has largely increased in recent years (Bradley 2006, 9). However, for some other people, these changes are not fully positive. For instance, it has been observed that despite the professionalization of fieldworkers, they are still among the lowest paid and career opportunities are still based in short-term contracts. Overall, this has created several problems for the development of an experienced workforce (Everill 2009, 3; Edgeworth 2011). A recent report of Historic England coincides with this diagnostic after the massive desertion produced by the economic recession in 2008 (Hook et al. 2016; Aitchison 2009).

The unification of digging and recording activities in one person was one of the most important consequences of the professionalization of fieldwork in Britain (Spence 1993, 25; Roskams 2001, 170). Nevertheless, this also implied the specialization of staff into fieldwork and post-excavation teams. Overall, this brought positive consequences as full-time diggers are able to accumulate a lot of experience compared with scholars that only spend a few weeks in the field every year (Everill 2010, 133). However, specialization also has constrained full-time fieldworkers from developing post-excavation skills (Everill 2009, 3–4). Nevertheless, one of the most frequent comments against professional diggers is their limited interpretative scope as '*the excavator's focus is very narrow and the level of interpretation minimal*' (Lucas 2001, 8). In other words, commercial diggers are characterized as recorders of ruins but not as true researchers of the past (Berggren and Hodder 2003, 424; Andrews,Barrett and Lewis 2000, 528). Similarly, many negative comments arose in the first decade of the commercial era due the adoption of a managerial culture that promotes standardization, preservation by record and a distinction between the recovery and the analysis of data (Shanks and McGuire 1996, 80; Adams and Brooke 1995, 96). These principles constitute the methodological core of British guidance such as MAP2 (English Heritage 1991) and MoRPHE (Historic England 2015). Hence, many critics consider that these principles explain the descriptive focus of fieldworkers (Everill 2009, 37; Lucas 2001, 14; Chadwick 1998; Hodder 1997).

In sum, despite commercial archaeology in Britain has many positive attributes like having a professional workforce, standardized methods and a good level of funding, many British archaeologists have very negative views about it. A group of them consider that competitive tendering is the main problem that has compromised the quality of projects and alienated the labour for fieldworkers (Thorpe 2012, 47). The strongest version of this view even subscribes the existence of a 'crisis in commercial archaeology' (Everill 2009, 3). For others, methodological aspects are the main factors that have affected the quality of research (Carver 2011, 10). The most radical version of this view comes from a group of theoretically minded scholars claiming that standardization of commercial practice restrains the interpretative components of fieldwork. In some extent, this group acknowledges the incidence of economic factors, but these are secondary (Hodder 1997; Andrews,Barrett and Lewis 2000). Then, as Richard Bradley has observed, British fieldworkers are the victims of their own success and they are very depressed about it (Bradley 2006, 9).

Reflexivity and the standardization of recording methods

One of the products of the rescue era was the development of single contexts recording (SCR) which occupies a privileged place in the history of British methods. However, this process has been interpreted ambivalently again. For some people, SCR represented a movement towards the professionalization and empowerment of excavators by giving them more responsibility in recording activities (Spence 1993, 25; Farid 2000, 24). For others, SCR also brought negative consequences because it became the most predominant method in Britain that produced an undesired level of standardization (Carver 2011, 20). Then, every site is reduced to a collection of context and every context to a list of attributes, which again explains the proliferation of description over interpretation (Lucas 2009, 232; Berggren and Hodder 2003, 423).

In this intellectual context, Reflexive archaeology has arisen as theoretical movement derived from post-processualism with the goal to develop an alternative fieldwork method that aims to solve the problems created by standardized practice (Hodder 2005, 644). Ian Hodder and his colleagues of the Catalhöyük Research Project are the most well-known promoters of a reflexive method and they are among the strongest advocates that standardization explains the proliferation of descriptive data in commercial records. For instance, they claim that context sheets usually provide layer descriptions without explaining how layers were defined in the first place. Hence, Hodder believes this creates a false view of layer descriptions as objective data which also fail to give account of uncertainties in interpretative practice (Hodder 2005, 651; Mickel 2015, 303). Moreover, reflexive archaeologists believe that standardized archives are difficult to understand and evaluate due the lack of enough contextual information (Hodder 1999, 95). These problems are summarized in the following fragment:

"*With the standardization of recording at all levels of analysis, we tend only to record what the forms (the recording system) require us to record. We tend, therefore, not to express worries, doubts, impressions, debates, inconsistencies. The problem with any text is that is difficult to understand when severed from its context of production"* (Hodder 1999, 31).

The solution of reflexive archaeologists consists in improving field methods, specially recording practices. For example, they reintroduced the use of excavation diaries alongside context sheets because they think that diaries have a less rigid format which facilitates to capture information absent in context sheets. Thereby, diaries will provide crucial contextual information to understand and evaluate primary data but at the same time they will improve the performance of fieldworkers by encouraging a more reflexive practice. These ideas are summarized in the fragment below:

> *"Reflexivity is also engendered by the diary writing and video filming, since these processes encourage those on the team to examine their own assumptions and provide contextual information about the excavation process, so others can look back and critically evaluate the claims that have been made"* (Hodder 2000a, 9).

The revival of diaries is based in previous characterization of context sheets as standardized formats and diaries as unstandardized (Barker 1982, 147; Hodder 2000a, 7; Farid 2000, 25). However, this is a rather simplistic view because systematicity always has been an important concern in laboratory and field diary recording (Mulligan 1996; Department of Physics University of Cambridge 2014; Stocking 1983). More recently, Shahina Farid noticed that reflexive diaries often contain irrelevant information (Farid 2015, 72), whereas Asa Berggren pointed out systematicity problems in reflexive diaries (Berggren and Nilsson 2014, 62) and Alison Mickel noticed redundancies in the content of diaries (Mickel 2015). These elements indicate that there are important theoretical problems in the methodological discourse of reflexivity as in one hand, it disqualifies standardization of procedures but on the other, reflexive methods cannot be implemented satisfactorily without it. Hence, these elements suggest that there are good reasons to question whether reflexive methods effectively have improved the quality of data. Particularly, there is an issue regarding the relevancy of such large amount of information almost in an attempt to capture every aspect of excavation work (Chadwick 2003, 103; Lucas 2012b, 72). This is rather ironic if one considers that early post-processualism heavily criticized rescue archaeology for similar reasons (Tilley 1989).

As mentioned above, one goal of the reflexive method is to improve the performance of fieldworkers and 're-empower' them for interpretative work because standardization and specialization apparently has restrained them to develop their interpretive skills (Berggren and Hodder 2003, 427). For this reason, Hodder has invited post-excavation specialist onsite for feedbacking additional data to fieldworkers and enrich their background knowledge to interpret the site. Likewise, reflexive method has facilitated the accessibility of databases onsite, so excavators can have access to previous reports and achieve a similar goal. Overall, reflexivity shows a recurrent pattern as it puts much trust in technology and tools for improving field practice. Moreover, reflexivity has been criticized recently because its attempts of empowerment usually leave aside issues of political representation, authorship and economical reward (Chadwick 2003, 103; Everill 2009; May 2012, 177). Besides, it has been observed that reflexive projects maintain traditional hierarchies between scholars and fieldworkers (Farid 2015, 76) that replicate master/servant relations between scientists and technicians (Shapin 1989). Finally, although reflexivity promotes a method that puts a lot of confidence in instruments, it still promotes an ideology in which technical work such as defining a feature is less valuable than interpreting an ancient culture (Shanks and McGuire 1996). This is very clear when scholar authorities claim that only when fieldworkers reach the highest levels of interpretations, they can be called archaeologists (Berggren and Hodder 2003; Andrews,Barrett and Lewis 2000). However, as Shapin explained long ago the value of intellectual over manual labour is historically grounded because science has been primarily defined as thinking rather than work (Shapin 1989, 561). Then, it shouldn't surprise that reflexivity prioritized deep thinking.

The feature system and the instrumentalist notion of recording methods

Before reflexivity, there was already a British tradition of methodological discussions of field methods. Many of those discussions are published in the proceedings of the Interpreting Stratigraphy conferences and a recurrent topic was identifying problems and limitations of single context recording (Steane et al. 1992). Some of those debates speak from the experience of having worked in landscape sites with not much stratigraphy (Clark 1992, 17) with alternative methods like the feature system, a recording strategy based in multiple sequences of excavation sheets hierarchically organized for the description of contexts, features and group of features (Thorpe 2012, 36). One of the goals of the feature system is encouraging higher levels of interpretation, particularly when working in deposits with remains that extend more horizontally that vertically, whereas SCR is more cautious about this, especially when working in deeply stratified sites (Hammer 2000, 143).

In recent years, Martin Carver has become a keen promoter of the feature system and a strong critic of SCR. He holds a less radical stance than reflexivity because he acknowledges that SCR might be useful for certain types of projects and sites, for instance when stratigraphy is deep and schedule is tight (Carver 2011, 28). But ultimately, he believes that a more complex interpretative system encourages more interpretation onsite (Carver 2009, 22). Carver has recovered the principles of the feature system for what he calls a multi-conceptual recording strategy (Carver 2009, 139). These thoughts are summarized in the following statement:

Stratigraphic excavators carry a "stratigraphic vocabulary" in their heads, and the more experienced they are the better chance they have of matching what they see to this vocabulary [...] In addition, to the description of the context, many excavators also carry more in-depth, *more interpretative inquiries -how did the context get there, what roles does it play in a human activity, which other context does it belong with? In this more comprehensive approach, grouping of contexts are defined onsite using another, higher-order stratigraphic unit -the feature, recorded on a feature card* (Carver 2009, 121).

Overall, Carver coincides with reflexivity as both believe that single context recording constrains interpretation, whereas a more complex method enables capturing a more interpretative data, especially higher-level claims. Gavin Lucas, resumes these views when that SCR is based in a geological language that favours the fragmentation of sites into layers (Lucas 2012b, 88, 2001, 152), whereas the feature system and excavation diaries encourages capturing higher-order data due its synthetic method (Lucas 2009, 232). The stratigraphic focus of SCR is obvious because it was designed to produce secure observations in complex stratified deposits (Hammer 2000, 143). Nevertheless, people from the Museum of London has argued that the system was designed to link descriptive and interpretative data which should reflect the rationale of interpretations (Roskams 2001, 171, 244; Spence 1993, 28, 40). In other words, the system was designed to give a reflexive testimony of fieldwork (Thorpe 2012, 41). Unfortunately, the view of SCR and commercial archaeology as purely descriptive work has gained more popularity over the years, even amongst its own practitioners (Watson 2019).

At the same time, the idea of fixing recording methods for improving interpretative work has gained more popularity too. For example, the Durham Archaeological Services create an alternative version of contexts sheets with important formal modification (Adams 2000, 94). Likewise, Adrian Chadwick, has presented a preliminary design of what he defines as 'more interpretive' format (Chadwick 2003, 108). Then, there is plenty evidence of a strong belief among many British archaeologists that a more complex recording strategy will improve interpretative practice and the quality of records. Some of these views acknowledge the relevance of additional factors like the skill of fieldworkers, but in general many of these views express great confidence in tools as if having more tools like reflexive diaries or 'better tools' like the feature method automatically will improve the performance of fieldworkers and the quality of data. Then promoting a strong instrumentalist view in which the quality of fieldwork mainly depends on the tools and methods being used. Thereby, the main problem that will be investigated in this thesis is an epistemological question that will be characterized as the problem of quality in excavation records or *In what extend an alternative recording system like the reflexive method and the feature system improve the quality of primary data?*

The problem of quality in excavation records.

This characterization of the problem of quality is based in the idea that single context recording, the feature system and reflexivity are alternative recording methods. However, this claim must be clarified because there is a degree of overlap. For instance, the reflexive and the feature system still make use of context sheets. Then, a good way to differentiate these recording methods is observing that SCR is based in a sequence of context sheets (Museum of London 1994), whereas the feature system includes additional sequences, for instance feature sheets, above the basic sequence of context sheets (Carver 2009, 140). Similarly, the reflexive methods as implemented by Hodder and his colleagues in Catalhöyük includes a sequence of excavation diaries on the top of different sequences of excavation sheets (Hodder 2005, 650). Based in the number of sequences of written records, SCR is the simplest method whereas the feature and reflexive system are more complex. In other words, my investigation aims to examine whether a recording system that includes more written information effectively produced higher-quality records.

Specifically, the problem of quality concentrates in the variability, credibility and relevance of information in written formats. For example, it aims to study whether a more complex recording strategy like the reflexive method and the feature system improve the variability of data, especially if they can give account of a wider range of interpretative information, and reflexive aspects like doubts and reconsiderations relate to excavation work. The issue of relevance aims to investigate the way in which additional information contributes to improve the credibility of primary data. Finally, it’s the issue of credibility of data which aims to evaluate whether additional formats enable the evaluation of records. These clarifications are important to define the universe of research but also to differentiate it from alternative ways to study quality in primary records, for example in graphic records (Bradley 1997; Lucas 2012b, 237).

The importance of the epistemic context in excavation work

In his doctoral investigation, Matt Edgeworth described the interpretative process of diggers that he characterized as the act of discovery (Edgeworth 2003). In general, the notion of the act discovery coincides with a lot with mainstream post-processual interpretive theory (Hodder 1991; Shanks and Tilley 1992). However, Edgeworth's publication explains in detail, the relevance of skill and experience in practical interpretations. Yet, his work makes less emphasis in recording activities, but the relation between skill and the recording work is very clear in the methodological discourse of Museum of London which insists that a method is not sufficient for having high quality records, it also requires a trained and experienced workforce (Spence 1993, 41; Roskams 2001, 170).

Sometimes, the merits of reflexive method have been questioned due the size of its budget and therefore doubting the applicability of reflexive methods in commercial scenarios (Farid 2000, 27; Chadwick 2003, 102). This argument is not very solid because many commercial projects also have large budgets than the average research project (Jennings 2011, 284). In consequence, although money is an important factor, it doesn't represent an important difference between academic and commercial contexts. Instead, there are other factors that define alternative epistemic contexts that seem to have stronger incidence in interpretive work. For example, Richard Bradley observes that academic projects often explore sites that have been investigated before whereas commercial projects often work in sites with no previous research (Jones 2013, 180). This contrast points out differences in the background knowledge of projects which most certainly will shape their interpretative results. Similarly, Gavin Lucas has noticed possible differences between short term projects in commercial archaeology and long-term research projects that facilitate the accumulation of background knowledge and revision of preliminary interpretations (Lucas 2012b, 235).

Finally, the survival of remains is another factor that had more relevancy in earlier interpretive theory. For example, Grahame Clarke acknowledged the incidence of the tools and skills of fieldworkers, but ultimately he considered that the fragmentariness of remains mainly determines what archaeologists interpret in the field (Clark 1957, 74). These views are connected as they portray interpretative work being shaped by the epistemic context of an investigation, which is defined by multiple factors like the skill of practitioners, the background knowledge of a site and the fragmentariness of remains.

Then, one of the main problems with instrumentalism in recent methodological discourse is being partially inconsistent with interpretative theory. For example, reflexive theory acknowledges the incidence of skill and background knowledge (Hodder 1999, 49) but in practice they put a lot of confidence in technology for improving excavation work (Berggren et al. 2015). For example, reflexivity doesn't pay much attention to necessary level of training and skills for using these diaries, video recording and so forth. Then, this promotes a view in which having more tools automatically will produce better results. Similarly, there are some inconsistencies in the methodological rhetoric espoused by Carver as in one hand he acknowledges the incidence of tools, terrain and the abilities of fieldworkers (Carver 2016, 42, 2011, 33) but on the other, he strongly believes that a more complex recording tool will suffice to improve the quality of records (Carver 2009, 121). Likewise, methodological discourses which supposedly demonstrates the success of reflexive method often ignores the epistemic context of reflexive experiments (Berggren et al. 2015; Hodder 2005; Andrews,Barrett and Lewis 2000). For instance, the most well-known reflexive project was mounted in the famous Neolithic site of Catalhöyük in Turkey where the survival of remains is very high and there is an important amount of previous investigation. In this way, one of the goals of my investigation is to consider what are the differences when investigating a site like Catalhöyük and a commercial site in London where evidence might be more fragmentary and background knowledge more restrained. Then, if one considers these elements, there are good reasons to question instrumentalist rhetoric which is not very consistent with interpretive theory.

In sum, the main goal of this investigation is to present an epistemic analysis that examines the relation between field methods, the epistemic context of investigations and the quality of records. Particularly, the analysis aims to investigate the incidence of various circumstantial factors the skill of fieldworkers, the background knowledge of projects and the character of remains in the use of field methods. Specifically, this investigation is grounded in a comparative analysis of three projects that implemented alternative recording methods, namely the reflexive system, the feature system and single context recording. This comparative strategy will provide a more robust judgment about the merits and problems of recording systems and interpretive practice in archaeological excavation.

Description of the thesis

To achieve this task, Chapter 1 presents a detailed revision of interpretive theory, putting special emphasis in the incidence of embodied cognition, techniques and observational circumstances in excavation work and how these aspects affect the relevancy and quality of data. Additionally, the chapter explores in detail the relation between background knowledge and the character of remains with interpretive work. One of flaws of recent methodological debate is being largely based in professional opinion, but poorly grounded in empirical evidence. Hence, Chapter 2 describes a methodology for a comparative study of three British recording systems based in a documental analysis of primary records. This chapter draws different procedures to evaluate the quality of records, especially the variability, credibility and relevance of information in context sheets, feature sheets and excavation diaries. Additionally, the strategy aims to measure the incidence of skill, background knowledge and the character of remains in field data.

Chapters 3-5 describe the analysis of three case studies. Chapter 3 describes the analysis of excavation diaries from Catalhöyük following the ideas of the reflexive method. Chapter 4 describes the results in a typical commercial site in London recorded with single context recording. Finally, Chapter 5 presents the results for the feature system as implemented in the explorations of the medieval site of Portmahomack in Scotland. Furthermore, these cases were selected for representing different epistemic circumstances, which include two academic projects and one commercial, two long term investigations and one short-term project. Likewise, case studies used slightly different types of workforce that explored different types of deposits. All the case studies were relatively well-funded and all of them published a monograph. The last feature was crucial for methodological reasons, but also aimed to minimize the incidence of economic factors.

Chapter 6 presents the discussion of results based in the comparative analysis of case studies. This evidence explains more clearly the incidence of epistemic circumstances in interpretive work. Furthermore, this chapter explains more clearly the methodological advantages and problems of recording strategies. The chapter makes general considerations regarding the role of feedback mechanism and strategies of representation using textual and visual testimonies. Finally, one of the characteristics of recent interpretative theory is being strongly based in an embodied perspective, hence one of the goals of this investigation is sketching an alternative view that reflects more clearly the coexistence of an individual and a collective dimension in interpretative practice. The investigation concludes returning to the starting point to revise to analyse the inconsistencies between methodological instrumentalism and interpretive theory. This time, however, describing other forms of instrumentalism grounded in technological improvements that promote a strong confidence and technology but at the same time aim to minimizes the performance of fieldworkers. Overall, this might be problematic in a discipline that lately has been characterized as an interpretative science.

Questions about recording methods are less popular today. Yet, this investigation might be relevant to examine opinions expressed in the last two decades which now have settled in. For example, a recent paper by John Barrett still claims that the answer for improving interpretative practices rests in enhancing methods (Barrett 2016, 137), equally most British archaeologists seem convinced about the sins of standardization, even commercial archaeologists (Watson 2019). Finally, there is now a new generation of reflexive archaeologists that still puts great confidence in tools and instruments for developing an epistemic virtue, namely reflexivity (Berggren 2015; Taylor et al. 2018). Nevertheless, it remains to be proved whether reflexivity has in fact improved interpretive practice.

Chapter 1 Interpretation in archaeological excavation

What an archaeologist finds when he sets the spade into the earth depends to some extent upon the methods he uses and the powers of personal observation he brings to bear; but the possibilities of any site are limited fundamentally by what has survived the passage of time. Grahame Clark, 1957, 74.

In recent times, archaeology is about interpretation for most British archaeologists. They interpret stratigraphy (Roskams 2000), landscapes (Tilley 2009), but they also interpret the Neolithic (Thomas 2013) and the axe trade (Bradley and Edmonds 1993). This small sample suggests two main interests. One for interpreting material remains and another for interpreting past events. This view of archaeological work largely corresponds with an theoretical developments in the last forty years. For example, theoretical studies in the eighties and the nineties mainly focused in the interpretation of the past (Wylie, 2002; Johnson, 2010) whereas the last two decades have seen a wider interest in the interpretative components of field practices (Edgeworth 2003; Lucas 2001; Hodder 1999). Specifically, interpretative issues coincided with the emergence of post-processualism when some archaeologists stablished an analogy between interpreting a text and investigating the past remarking the importance of context and subjective elements. This work was primordial to trace a border with mechanical practices aiming for objectivity and veracity (Shanks and Hodder 1994; Hodder 1986). However, this characterization of interpretation is not fully accurate. For example, the interpretation of ancient texts is often regulated by rules. Likewise, interpreting the law and documental evidence is also regulated by rules that aim to reach veracity (Ginzburg 1999). Performances like playing music or interpreting a play are other analogies that has been used to explain the interpretative nature of archaeology (Tilley, 1989). Again, these exercises are regulated by rules and this is more evident in music. Otherwise, what would be of an orchestra where each player assigned different meanings to the notes? However, not everything is determined by rules. For example, Shakespeare’s scripts include sequences of dialogues and scenes, but they are more ambiguous about how an actor should declare his lines. Hence, these aspects are more open to interpretation and probably will be derived by contextual information on texts and background knowledge about previous performances.

In short, many interpretative exercises depend on following certain rules, but they are not reduced to a systematic application of procedures. The subjective aspects of performance and the context of words and actions are important too. Hence, post-processual discourse fails to see the relation between mechanistic and subjective elements in interpretation. This shortcoming creates some tensions between post-processual interpretative theory and methodological discourse which often praise mechanization and systematicity. For example, post-processual thinkers often claim that standardized practices generally reinforce objectivity in the detriment of subjectivity (Shanks and McGuire 1996; Shanks and Tilley 1992). To solve this dilemma, reflexive archaeology has attempted to embrace subjectivity incorporating what they consider less standardized practices like diary recording (Hodder 1997). In this process, nevertheless, reflexive method has put greater confidence in technology which raises important inconsistencies with interpretative theory.

In general, interpretative theory suggest that what archaeologists interpret depends on their epistemic circumstances defined by multiple factors like their skill and experience (Edgeworth, 2003). Long ago, Grahame Clark framed something similar as the epigraph in this chapter proves. Likewise, Ian Hodder seems to agree with this view: "*How much archaeologists can interpret depend on the richness of their data networks and on their knowledge and abilities*" (Hodder, 1986, 176). However, as explained before methodological instrumentalism and interpretative theory don't match very well due a strong confidence in tool and a minimization of other epistemic factors. Besides, another important inconsistency in post-processual discourse later incorporated into reflexivity is the conceptualization of interpretation and mechanization as opposites, however as it will be explained this is an untenable stance in archaeology.

The goal of this chapter consists in describing in more detail the inconsistencies between interpretative theory and methodological discourse. At the same time, this chapter aims to provide a more satisfactory explanation of how standardized and subjective practice are integrated in technical work, especially excavation and recording. Hence, this chapter begins by providing a general distinction between the interpretation of remains and the interpretation of the past. This distinction is important to capture different forms of gaining knowledge: direct experience and inference (1.1). Once this ground has been laid, the following section considers the role of skill in embodied cognition (1.2). Afterwards there is a short explanation of the relation between skill, technique and standardization in embodied perception to produce reliable observations and reports (1.3). By its own nature, methodological discourse stresses the mechanic elements of excavation nevertheless this is also a learning process that constantly changes the views of fieldworkers. This dual nature is best captured by the notion of the act of discovery which is also very successful to reconcile the mechanic and subjective elements of excavation (1.4). Each one of these sections also compares how these aspects have been considered within methodological discourses of the Museum of London and Reflexive archaeology.

The second goal of this chapter consists in explaining the incidence of additional factors that define the epistemic context of an investigation. One of these sections explains the role of background knowledge and expertise (1.5). Later, there is a short explanation of how the character of deposits and remains influences interpretative and recording practices (1.6). At this point, some preliminary arguments are drawn to explain how these aspects have not been considered within reflexive rhetoric. Afterwards, there is a brief explanation of how excavation strategies influence interpretative work (1.7). One of the characteristics of recent interpretative theory is defining excavation work as a subjective or embodied experience, nevertheless this view can be enriched because fieldworkers also act like an epistemic community. Therefore, Section 1.8 draws some ideas about a collective dimension of interpretive work.

1.1 A conceptual distinction

According to Gavin Lucas, interpretation and explanation were indistinct terms to characterize the investigation of the past before the days of New Archaeology. Afterwards, the notion of interpretation became suspicious due its association with subjective opinion (Lucas 2014, 4015). Evidently, the revival of interpretation came with post-processualism and the main conflict with processualism laid in the type of explanations to be produced and its credibility (Trigger 2006; Wylie 2002). Despite such differences, the methodological principle of middle range theory and interpretative archaeology are practically the same (Kosso 2001). Following the textual analogy, Hodder suggested that "*interpretation is translation. It involves the archaeologist acting as interpreter between past and present.*" (Hodder 1991, 15), whereas middle-range theory occasionally was described in similar terms:

> *The challenge that archaeology offers, then, is to take contemporary observations of static material things and quite literally,* ***translate them*** *into statements about the dynamics of the past ways of life* (Binford, 1983, 20).

However, as Lucas explains, no explicit discussion of the notion of interpretation arose until the second stage of post-processualism when it changed from *contextual* into *interpretative archaeology* (Lucas 2014, 4016). Then, hermeneutic influences became more evident and interpretive discussion developed in two fronts. First, Hodder described a hermeneutic model for archaeological inference characterized as a process of fitting multiple data into a coherent explanation. However, as more data is collected, preliminary explanations are revised and corrected, thereby producing a recursive loop between data and interpretation (Hodder 1999, 64). In general, the hermeneutic model is an accurate description of the process of hypothesis formulation, but the recursive mechanism of empirical science had been already studied by models of non-deductive inference, particularly with the investigations about the process of scientific discovery inspired in the philosophy of Charles S. Peirce (Aliseda 2004). Moreover, an additional aim of the post-processual hermeneutic model was stressing the distinction between archaeology and experimental science, by showing that archaeology doesn't work by testing hypothesis. Although, this claim is true, it's based in a poor understanding of experimental science as if laboratory work only included demonstrative experiments, without considering its multiple stages such as stabilizing and debugging experiments (Hacking 1975, 37).

The second front of interpretive archaeology was a reconceptualization of data production and technical work as interpretive exercises. In broad terms, this argument claims that observations, description and classification are interpretations because data are not given but instead they are constructed in relation to a series preunderstandings and a selection of attributes (Shanks and Hodder 1994, 7). In general, this was common ground among interpretive archaeologists, although there were subtler differences among them. For instance, Tilley considered that interpretation is an epistemic situation that only occurs when someone is puzzled because something is not obvious. For example, he said, if you see an animal in the distance without being sure whether it is a dog or a badge, then this is an interpretative judgment (Tilley 1993, 2–3).

On the contrary, Julian Thomas went a step ahead and claimed that every act of experience is interpretative because perception is selective and is organized by our background knowledge (Thomas 2004, 30). In general, this arguments replicates the idea of observation being 'theory-laden' (Hanson 1958, 19). Finally, Hodder encapsulated many of these thought into a single notion when he claimed that "*interpretative judgments are present in all areas of archaeology, even down to making catalogues, doing laboratory experiments and excavating features in the field*" (Hodder 1999, 67). Thus, he included the notion of selective and theory laden observation to describe the process of layer definition when he claimed, "*there is no fixed entities of bounded earth that can be recorded independent of interpretation*" (Hodder 1999, 86). Besides, he recycled the notion of interpretative uncertainty when he explained that "*in practice we have to excavate without knowing what we are excavating*" (Hodder 1999, 92). All these elements are summarized in the notion of 'interpretation at trowel's edge' that links technical work with interpretative activity.

Nowadays, there is great awareness of these two broad areas of archaeological interpretation that could be labelled as *interpretation at trowel's edge* and *reading the past* to characterize the interpretation of remains and the interpretation of the past, respectively. Likewise, there is a clearer idea of the recursive relation between evidence and interpretation. However, post-processualists questioned the distinction because both aspects are closely related, for example noticing the evidence to interpret the function of a feature might be triggered by previous knowledge (Hodder 1999, 102; Thomas 2004, 30). However, the distinction is valid because it differentiates the observability of archaeological remains (evidence) and the unobservability of the past (interpretation). In other words, the distinction indicates two ways of gaining knowledge. A direct empirical experience of remains and an indirect knowledge of the past by inference. To be clearer, archaeologists can see remains, but never experience the past[1] (Bloch 1954, 40; Johnson 2010, 12).

The inferential knowledge of the past is a distinctive characteristic of historical disciplines, including natural sciences such as geology and cosmology. But the indirect knowledge of unobservable phenomena by observable evidence traces is a common feature of empirical sciences

[1] Some phenomenological doctrines redefine the past/present distinction (Embree 1992). But post- processualism has not provided a clear distinction between phenomenological and empirical experiences (Johnson 2006, 125).

and this includes the study of microparticles, microorganism and far away phenomena (Turner 2007, 23). This inferential relation from observable things to unobservable phenomena constitutes the basic core evidential relations (Hacking 1975, 34; Woodward 2000, 165). To give a simple example, a footprint is evidence of the someone been there. Nevertheless, some philosophers have explained that the distinction between observable and unobservable phenomena constantly changes due instrumental progress which enables to see more things like micro-organism or distant objects. This is very clear in archaeology as instruments frequently improve the observation of remains and traces, however in the case of past phenomena the frontier is stronger, because past events will not become visible despite instrumental progress.

More recently, some people has remarked the practical component when interpreting evidence against the theoretical nature of interpreting the past (Edgeworth 2003, 6). Yet, this has raised some strange responses claiming that theory is also a practice due the use of conceptual tools and rhetoric skills (Tomášková 2006, 165–166; Shanks and Tilley 1992, 116). However, this view is mistaken because theory is not about writing nor giving speeches, but certainly theoreticians use these tools. Theory is about making general statements, which are completely different to practical work consisting in using general principles for solving specific situations. Hence, Newton's Principia are not theoretical for being written rather than being demonstrated by example (for instance with a falling rock), the Principia are theoretical because they explain different kinds of phenomena (falling objects, tide changes and planet orbits).

Finally, there is an additional reason to preserve the distinction between evidence and interpretation as both are related to different epistemic problems. To count as evidence, an observation must be reliable or successful to control noise and reduce errors (Woodward 1989, 394). On the contrary, the credibility of interpretation depends on the robustness of evidence. Then, some interpretations might be barely suggested by evidence whereas others might be strongly indicated by multiple strands of evidence (Kosso, 2001). Considering these criteria, explanatory models such as the hermeneutic model only gives account of the formulation of a plausible explanation that ultimately could be flaw if observations are erroneous (Aliseda 2005). Considering the above, I follow the conventional distinction between evidence and interpretation. Even-though, in a more general sense both are result of interpretative exercises.

1.2 Observation and skill

For traditional empiricism, observation is a process by which senses receive inputs (or sense-data) and constitutes the foundation of knowledge. Therefore, observation produces accumulation of experiences that subsequently allows to formulate theories (Ladyman 2002, 33). This view dominated until the first half of the twentieth century, when some positivist philosophers began to question the foundational role of experience (Popper 1972, 46). One of the most well-known critiques is the notion of observation as theory-laden proposed by Norwood Hanson. To introduce his argument, Hanson imagined Tycho Brahe and Johannes Kepler looking at the sun, then he argued that Kepler would see a moving star and Tycho would see a fixed body, due their geocentric and heliocentric beliefs. Hence, Hanson explained that although both astronomers had the same visual stimuli (or sense-data), they saw different things because observation is visual sensation organized by background knowledge. Thereby, the empiricist mistake had been reducing observation to physical stimuli (Hanson 1958, 5–19). In consequence, perception (or sense-data) could no longer be considered as the ultimate ground of knowledge.

More generally, the thesis of the theory-laden observation claims that perception involves conceptual interpretation because visual stimuli is organized by background knowledge such as concepts, beliefs and theories. Another typical example of seeing different due different background knowledge are the perceptions of the layman and the specialist, for example a trainee in the field would see rocks and earth, whereas a trained fieldworker would see layers and features (Edgeworth 2003, 21; Lucas 2012b, 16). Nevertheless, this example suggests that people can learn to observe in new ways.

More recently, Hans Radder has argued that the notion of theory-laden observation is incomplete because it fails to consider the actions of observers. For example, if someone is having difficulties seeing a distant object, he can get closer. Likewise, if there is not enough light, people can use a torch to improve visibility. Thereby, Radder claims that observation includes *material realization*, namely the actions and preparations that enable seeing. Specifically, he considers body movements and adjustment to observational conditions (Radder 2006, 31). In consequence, the theory of material realization coincides with views of embodied cognition and phenomenology that stress the role of agents as crucial elements for perception (Radder 2006, 74).

Similarly, some philosophers previously had to remark the role of instruments for observing things because some interpretations of the theory-laden doctrine suggested that new concepts made things visible (Hacking 1983, 179). For example, it would be like assuming that before the definition of the notion of the cut, archaeologist couldn't see pits and ditches. Instead, what the stratigraphic concept does is reformulate the ways that archaeologist perceive such things, not only as holes in the ground but as stratigraphic events. Instead, it is fair to say that techniques enable or obscure the observation of cuts. For example, in Mesoamerican archaeology there is a recurrent practice of digging arbitrary levels and many photographic records show burials over rectangular soil beds. Similarly, Chris Evans has explained how technical changes introduced in Britain by Gerard Bersu facilitated the detection of post-hole structures (Evans 1989; Lucas 2012b, 216). In short, instruments improve visibility and theories help to conceptualize what is being perceived. This doesn't mean that observers can avoid conceptual interpretation but being a good observer does not require being aware of all the underlying theory (Radder 2006, 79). Mathew Johnson claims something similar when he asserts that diggers constantly use theories of soil to detect stratigraphic change, even if they are not fully aware (Johnson 2010, 22).

More recently, Martin Carver has stressed the role of instruments for improving visibility in archaeological deposits, especially to explain how different tools such as trowels, shovels and laboratory tools allow to observe different types of entities with different levels of precision (Carver 2011, 37). However, Radder explains that material realization doesn't imply that every observation will be successful because one might fail to use an instrument or prepare a sample (Radder 2006, 79). At this point, one must consider the role of techniques because these procedures have a crucial function to create adequate observational conditions to prevent misleading observations derived from working with uncalibrated instruments or contaminated samples (Woodward 1989, 396). In short, techniques regulate the use of instruments for producing reliable observations.

Similarly, Ian Hacking highlighted the role of skill and training for using instruments to create reliable observations (Hacking 1983, 168). This doesn't mean that skilled observers are infallible, but experience teaches them out to identify when something goes wrong. In this way, human error is a general source of unreliable observations, but one can set a distinction between errors by lack of training and technical accidents (Allchin 2001, 6; Hon 1995, 9). Altogether, these notions illustrate that actions and circumstances can affect positively or negatively the quality of observations. For example, Grayson and Meltzer questioned the validity of some explanation regarding the extinction of Mammoths due predatory activities by Clovis hunters, because evidence of ‘killing-sites’ was suspicious due different problems with recovery procedures (Grayson and Meltzer 2002).

Nevertheless, one must differentiate cases of unreliable observations by error with cases when the confidence of an instrument or a technique is questioned in the first place. One of the clearest examples is the dispute about arbitrary methods which are questioned due the destruction of natural deposits and merging of artefacts. Thereby, producing deceiving observations that create unreal layers and assemblages (Harris 1989, 20). However, Martin Carver explains that arbitrary techniques create flat surfaces with numerous truncated layers that are recorded in plan and complemented with section records. In this way, arbitrary techniques produce more controlled observations by cross-checking plans and sections (Carver 2009, 117). Additionally, he remarks that many projects follow arbitrary methods when diggers are not sufficiently trained to follow natural layers. Hence, arbitrary techniques regulate the action of technicians by keeping flat surfaces and reducing the possibilities of making thigs up (Carver 2011, 21).

This connection between techniques and skill for the production of reliable observation is a recurrent issue since the emergence of modern archaeological fieldwork (Carver 2012, 17; Lucas 2012b, 37). For example, most excavation handbooks usually praise a set of technical procedures and the abilities of fieldworkers (Barker 1982, 12). Similarly, single-context recording is grounded in the confidence of a set of procedures and the training of fieldworkers that enables them to produce trustable information. For these reasons, it’s very clear for single-context recorders that having a method is not enough (Roskams 2001, 170). In contrasts, post-processual interpretative theory generally acknowledges the relevance of skill in technical work (Shanks and McGuire 1996; Berggren and Hodder 2003). However, when describing reflexive methods instruments are the main source of trust whereas skill and training seem irrelevant (Hodder 2005; Berggren et al. 2015; Mickel 2015). In that sense, it’s ironic that although Hodder frequently questions the skills of commercial diggers, they were the main workforce of his project in Catalhöyük. Fortunately, the relevance of skill was not unnoticed for everyone. For example, Shahina Farid suggested in one of the earliest drafts of reflexive method that: “*To achieve a ‘reflexive excavation methodology’ …a team of more fully proficient excavators should be employed who know how and when to record information*” (Farid 2000, 28). Unlike Hodder, she is aware of the relevance of skill for running a reflexive practice. Hence, in a subsequent retrospective exam, she observed that only experienced fieldworkers were able to record relevant data in reflexive diaries whereas the less experienced diggers usually captured meaningless observations (Farid 2015, 72). More recently, reflexive archaeologists have gained awareness of the importance of training and skill for recording activities, particularly after the introduction of tablets (Taylor et al. 2018). However, this judgment has not extended to further domains of reflexive tools like diaries and videos. Then, an important question to consider is whether making a reflexive record is just a matter of using a set of tools, independently if one's is not very good observer or whether being able to produce a reflexive record demands a level of training and skill.

1.3 Techniques and standardization

The previous section has sketched how technical skill contributes to produce reliable observations. Then, this section aims to explain how techniques define procedures for controlling observational conditions and the performance of fieldworkers for this purpose. Overall, this is fundamental to understand the role of standardization and systematicity in interpretative practice. For example, one of the commonest misconceptions about fieldwork is that for being outdoors, it is completely different to laboratory work which is the paradigmatic place of a controlled environment (Livingstone 2003, 42–45; Heggie 2014). This distinction might seem evident as excavations usually take place in the countryside or in construction sites. However, if Barker drew a separation between excavation and the laboratory, he rightly stablished a link with surgery which relates the field with the operating theatre, two forms of controlled environment for non-demonstrative experiments (Barker 1982, 12). One of the first things to notice in a site of archaeological dissection is the way to organize space in different areas. For example, excavation (trenches), finds and sample processing, storage (offices and cabins) and discard (spoil-heap). Moreover, these areas are connected in two processes. One for recovering and storing materials and another for discarding spoil. Similarly, pathways and accesses regulate movement across the site. Not to mention the grid and benchmarks which are additional ways of organizing the space for recording of evidence (Edgeworth, 2003, 19). Overall, excavation sites are cultural setting as any place of research, but instead of bringing the body to the operating theatre, a scenery is constructed around it (Figure 1.1).

Figure 1.1. A site of archaeological dissection. Burdale, Yorkshire. (Copyright Julian Richards CC BY 4.0). https://doi.org/10.5284/1021540

Figure 1.2. A site of archaeological dissection. North area excavation. Catalhöyük, Turkey. (Copyright Catalhöyük Research Project. CC BY-NC-SA 2.0).

More importantly, many site preparations contribute to control the observational conditions. For instance, shelters are installed to control rain, sun and wind. However, if a site is to dry, watering will improve visibility (Barker 1982, 104–105; Roskams 2001, 105). Evidently, site preparations will vary from case to case, but in general they contribute to control noise and reduce accidents. As Martin Carver says, it means protecting trench edges and pathways, so that people do not tread on clean surfaces and by the same token guaranteeing the safety of fieldworkers (Carver 2009, 123). Thus, even if archaeological sites are not fully controlled environments, the difference is more a matter of degree (Figure 1.2). Although, the frontier occasionally disappears when an excavation block is lifted to be dissected in the laboratory.

Technical procedures also contribute to control observational conditions by regulating the actions of fieldworkers. For example, layers must be removed one by one and in the inverse order of deposition. Likewise, technical rules can define the way to perform small scale actions like trowelling and even the way that fieldworkers move their bodies (Figure 1.5).

When actually removing a unit, one must always work systematically backwards, pulling the excavated material onto the unexcavated part at one's knees rather than pushing it away onto the already exposed and clean underlying strata (Roskams, 2001, 227).

Handbooks are full of the dos and the don'ts regulating technical work. Thus, Philip Rahtz recommends: "*The trowel must be used fast and accurately to remove soil and clean surfaces and there must be not 'follow-through', but rather an ability to brake hard*" (Rahtz 1974a, 276), which is a crucial to avoid merging material from different layers and therefore creating unreal boundaries. Technical rules will vary for different instruments such as the trowel and the mattock. Likewise, technical rules will vary if one is working in a test-pit as this requires moving down or if one is digging within an open area that requires moving sideways. But in general, technical standards aim to improve the reliability of observations by controlling environmental factors and the actions of fieldworkers. And for this reason, one must define archaeological excavation as the disturbance of the ground in controlled conditions (Roskams 2001, 1).

Hence, the development of technical procedures is rooted in the issue of the role of archaeologist as producer of primary data (Lucas 2012b, 37). However, this preoccupation can be inserted into a wider context in which the presence of an observer can risk knowledge, then one of the main aims of objective science is to develop mechanisms to regulate their behaviour until they become a sort of registering machines (Daston and Galison 2007, 38). However, as Edgeworth has explained, although these behaviours seem natural, all of them are learned. A result of such automatization of practice is the development of a technical virtue grounded in a discipline of the self and technical skill until one becomes a digging machine that systematically follows the standard of good practice. Frequently, this practical virtue becomes invisible, but its presence becomes evident when one compares the work of a novice and an experienced digger because the novice generally obscures rather than highlight significant patterns (Edgeworth, 2003, 35). Thereby, technical virtue initially arises from mechanization which also aims to regulate the observations and skills of different individuals (Roskams 2001, 170).

Technical standards also regulate the movement of archaeologists during recording activities. For example, using the planning framework requires moving around to produce accurate data and avoid misrepresentation. Yet, this is a complex issue because many recording systems include some level of systematic distortion. In that sense, accuracy is always relative to a set of representation rules. In the case of archaeological planning, this implies avoiding a perspectival view and favour a sort of ‘no-where’ view to place objects within a cartesian space (Lucas 2012b, 242; Edgeworth 2003, 99). Similarly, technical standards also regulate the process of filling up context sheets, for example layers must be described during excavation, and not afterwards when a unit has been removed. The purpose of this is to guarantee accuracy, by avoiding relying on memory (Roskams 2001, 170–171; Spence 1993, 32). Recording standards also indicates what type of information must be recorded and how this should be described (Museum of London 1994). This is most evident in layer descriptions and stratigraphic relations where descriptive terminology and representation rules are more standardized. Nevertheless, reporting interpretations is also regulated by rules as these must have some form of justification explaining what evidence supports it. In this way, mechanization and standardization work together to guarantying the quality of data. Then, the quality of primary data can be compromised by not following these rules. For example, Craig Spence has observed that one of the commonest mistakes in context sheets are unjustified interpretations (Spence 1993, 32).

Figure 1.3. The discipline of recording primary testimonies onsite. Burdale, Yorkshire. (Copyright Julian Richards CC BY 4.0). https://doi.org/10.5284/1021540

Paradoxically, mechanization and standardization have become suspicious in reflexive discourse. For example, Hodder says: "*With the standardization of recording at all levels of analysis, we tend only to record what the form (the recording system) require us to record*" (Hodder 1999, 31). Similarly, other archaeologists have claimed that standardization produces an uniformization of records (Chadwick 2003, 110; Lucas 2009, 232). Evidently, one cannot deny that standardization promotes a level of uniformization, otherwise it wouldn't be achieving its function, However, what is at stake is whether standardization and mechanization restrain the observational powers of fieldworkers to notice crucial attribute not listed in predetermined formats, which therefore could compromise the quality of records. Thereby, we reach an inflexion point, because in the first-place automatization contributes to the development of observational skills. But on the other hand, the same process seems to constrain the development of further observational powers to notice other things beyond the domain of formats.

For this reason, reflexive diaries are conceived as a useful tool to record a wider range of data due its less standardized format. Yet, there are some important evidence against this idea. For instance, fieldworkers were not very systematic with diary recording in the early years of the CRP, hence a stronger policy was required, and fieldworkers were instructed to record at least two weekly entries. Likewise, many participants didn't have much idea of the goals of diary recording and this required the definition of clearer guidelines (Berggren and Nilsson 2014, 62;

Berggren et al. 2015, 436–437). Evidently, these are important things to consider as inconsistency affects the quality of primary records. Similarly, the recent introduction of the tablets for recording activities had some negative effects as some fieldworkers adopted the bad habit of recording primary data outside the trenches. Hence, protocols had to be reinforced to maintain systematicity "*planning outside the trench is prohibited without good reason*" (Taylor et al. 2018). In this way, there is an important paradox in reflexive thought as in one hand, it rejects mechanization and standardization of practices, but on the other it's impossible to implement reflexive methods without them. This contradiction has softened in recent years as reflexive archaeologists have begun to become aware of the importance of systematicity (Berggren 2015, 6215). However, as explained before, this contradiction probably arose due a misinformed notion about the relevance of recording rules in diary recording for empirical sciences (Stocking 1983; Department of Physics University of Cambridge 2014). Still, this is not the whole story because excavation is not a completely mechanistic process, it's also a learning process and primary testimonies should reflect this in some extent.

1.4 The act of discovery and the expertise of fieldworkers

One of the most distinctive characteristics of archaeological excavation is that one must unearth things to observe something. For example, clear the soil to observe a layer, remove a fill to observe a cut or clean overlying deposits to see structures. Hence, this observational process enabled by the discovery of remains is different to everyday observation. This is an important precision because in the original formulation of the theory-laden observation, Hanson considered there could be exceptional cases of 'interpretative observation' when it takes time to observe something. For instance, if looking something at the distance in a foggy morning (Hanson 1958, 10), which is very similar with Tilley's account of interpretation in uncertain situations (Tilley 1993, 2). Hence, one must take account that a lot of uncertainty in archaeological excavation is produced by the general observational conditions in which an unobservable thing gradually becomes visible. Matt Edgeworth has provided one of the best accounts of this process that he defines as the act of discovery which describes an epistemic situation defined by the constant movement of spoil creating a threshold between the visible and the invisible that constantly redefines what is known and unknown (Edgeworth 2012, 80). However, another important characteristic of the act of discovery is an engaged observation that constitutes a multisensorial bodily experience:

"the cut is experienced by the digger as much more than just a visual phenomenon. It's 'felt' as difference in texture, with the trowel becoming an extension of the body for perception as well as action. Sometimes the cut has a sound: the noise of the scraping action of the trowel changes as it meets different soil on the side of the feature. Sometimes the new surface even has its own smell" (Edgeworth 2012, 79).

Still, the act of discovery also includes some forms of disengaged observations, for example when diggers stop trowelling and stand out for examining remains from the distance. Hence, the act of discovery includes shifts between an engaged and disengaged observations (Edgeworth 2003, 34). Another important characteristic of the act of discovery is an intrinsic fluidity because the exposed areas provide information that generates expectations about the unexposed remains. Nonetheless, predictions not always result as expected and diggers must readapt their strategy to cope with unexpected evidence that contradicts preconceived ideas (Edgeworth 2003, 33, 2012, 78–79). This indicates that during the act of discovery and the observational process, diggers make a lot of conjectures regarding the nature of remains. And this challenges the previous distinction between the direct knowledge of remains and the inferential knowledge of the past. However, the distinction can be maintained because inferences within the process of discovery are generally predictive whereas inferences for interpreting the past are retrodictive.

Overall, the act of discovery can be characterized as a progressive learning process about the character of remains that guides technical actions. For example, when excavating a fill, this should be dug horizontally if containing a burial, or excavated in section if having a posthole. Quite often this decision will be based in available evidence as the shape of the cut, the colour of the fill and previous experience with similar features. However, this information might be insufficient to predict what is underneath and fieldworkers will have to correct their technical decisions in the light of new evidence (Carver 2009, 131). For these reasons, technical work cannot be reduced to the mechanic aspects. Often, excavation is compared with other classes of technical exercises like playing music, but the metaphor has its limits because playing an instrument can be made blindfolded. On the other hand, excavation always requires constant monitoring for updating information (Edgeworth 2003, 54). In this way, the digging machine constantly interrogates his own perceptions and reconsiders his own decisions.

> *Routinely, as excavator we question the evidence of our senses, we rework our assumptions, we deploy our own contextually tailored versions of Cartesian systematic doubt, doubt which we exercise in those fractions of seconds between, and during the movement of the hand and the impression of the sense. We modify our individual strategies as we try to reconcile often conflicting impressions of what we think, we feel, and we experience during the process of excavating* (Thorpe 2012, 34).

The characterization of field interpretation as a learning process coincides with some aspects of what Daston and Galison have called trained judgement or a set of practices for the production and/or interpretation of data grounded in the expertise of practitioners (Daston and Galison 2007, 311). The trained expert has a basic set of technical skills for the interpretation of images created by mechanic procedures like photographs and encephalographs. However, the most valuable asset of the expert is his ability to recognize subtleties among series of relatively similar phenomena (Daston and Galison 2007, 344). In other words, trained judgment is skill enriched by experience to notice crucial aspects, impossible to detect by purely mechanical means. A similar idea is described by the notion of *connoisseurship* which also implies the refinement of observational skills after having examined multiple cases of similar phenomena that develops the ability to notice marginal but crucial evidence (Ginzburg 1989, 97). In this way, the trained expert develops a professional eye that integrates the mechanisms of objective practice and the subjective view developed by the accumulation of experience.

As explained before, the basic elements of archaeological excavation are grounded in a mechanical practice that contributes for the development of technical skills. Yet, at the same time practice contributes for an accumulation of subjective experience that enlarges the background knowledge of fieldworkers. For instance, how a roman feature look like or how a prehistoric structure looks like, and this empirical background supports technical decisions of what to do in a specific situation. This intimate relation between technical skill and trained judgment in excavation might not be evident, because excavation handbooks usually stress mechanical elements. But this doesn't mean that fieldworkers don't acknowledge the importance of expertise (Barker 1982, 12; Harris 1989, 53). Hence, it's rather contradictory when some theoreticians ask for handbooks that include guidelines for specific situations like how to excavate a Mesolithic floor (Olsen et al. 2012, 63).

Finally, the result of the act of discovery is an epistemic change in which uncertainty disappears by making things visible. For example, by exposing a structure (Figure 1.4) or exposing a cut after removing the fills. Yet, when the exploratory activity continues, for instance, searching for the next deposit, the act of discovery restarts and uncertainty emerges once more. Moreover, Edgeworth notices some similarities between following a cut and tracking an animal (Edgeworth 2012, 78) because both process describe a process of following an unobservable creature that eventually becomes visible. This situation is nicely illustrated by the following example that aims to explain the difference between observing something and having evidence of something.

> *The situation in which I would properly be said to have evidence for the statement that some animal is a pig is that, for example, in which the beast itself is not actually on view, but a can see plenty of pig-marks on the ground outside its retreat. If I find a few buckets of pig-foo, that's a bit more evidence, and the noises and the smell may provide better evidence still. But if the animal then emerges and stands there plainly in view, there is no longer any question of collecting evidence; its coming into view doesn't provide me with more evidence that's a pig. I can now just see that it is* (Austin 1962, 115).

However, this explanation is not completely satisfactory because this form of observational process is not always the case. It's also frequent that soil layers are progressively observed and removed without ever having an open view of those deposits or as Martin Carver says: "*An archaeological deposit is a three-dimensional artefact, only seen once and never seem whole*" (Carver 2009, 123). This is what Richard Bradley describes as one of the constrains of excavation because to observe something you must destroy was is blocking the view, but you can observe it only for a limited amount of time (Bradley 2003, 155) (Figure 1.5 and Figure 2.3).

Figure 1.4. Exposed buildings in the North Area, Catalhöyük. (Copyright. Catalhöyük Research Project. CC BY-NC-SA 2.0).

Figure 1.5. Two excavators removing a fill to define a cut. Burdale, Yorkshire. (Copyright. Julian Richards CC BY 4.0). https://doi.org/10.5284/1021540

In sum, the act of discovery is a progressive exploration of remains with two possible paths. In some cases, excavation will expose remains to plain sight, but in other cases, observations will be partial and progressive. These possibilities are relevant for their implications for recording activities. A common idea is that many site records produce a ‘no-where views’ like plans where perspective and subjectivity is eliminated (Lucas 2012b, 242; Edgeworth 2003, 99). This coincides with notions from the phenomenology of perception by remarking that embodied observers always things from a specific point of view and then they must move an object or move around to observe all its faces (Radder 2006, 29). Hence, an ‘absolute view’ of an object should condense multiple gazes into one (Merleau-Ponty 2011, 72).

Similarly, layer descriptions have been characterized as objective or abstract accounts of contexts which are disassociated from the process of discovery. For example, Hodder believes this a consequence of standardization (Hodder 1999, 95). Similarly, Edgeworth coincides with this idea as he characterizes the act of discovery as fluid whereas he describes recording as static because soil descriptions are suspended of time and fragmented into multiple attributes (Edgeworth 2003, 96). However, the process of layer description is characterized slightly different by Steve Roskams and Craig Spence as both suggest that context sheets should be progressively filled up. For example, they recommend making a preliminary description that later should be updated, particularly to include attributes not observed in the surface. This is a crucial aspect, because this process will enable to reconsider preliminary interpretations in the light of new observations (Roskams 2001, 171; Spence 1993, 28).

These opposed views about context recording are one of the main causes of conflict regarding the nature of standardized methods. In one hand, reflexive archaeologist believe that context sheets erase every trace of the act of discovery, particularly the fluidity and the rationale of interpretative decisions. Then, one of the purposes of diaries is to give testimony of these aspects (Mickel 2015, 303). On the other hand, Roskams and Spence suggest that despite context sheets will not provide a testimony of the act of discovery, descriptive and interpretive testimonies will be parallel to this process and therefore, records will provide an indirect testimony of the fluidity of this process. Hence, this controversy takes us back to the main issue or whether standardized formats reduce recording work into mechanistic practice as critics say or whether these formats can give place to a reflexive testimony of interpretative work. Nevertheless, an important problem with this discussion is ignoring the sequential nature of recording systems. In other words, single context recording is not the context sheet, but a sequence of related documents. Hence, this debate must consider that a sequence of context sheets is most likely organized in a series of conditional events that progressively reflect the learning and reflexive process of excavation. Although, another possibility is that sequences might be constituted by series of independent events and this would constrain its capacity to give account of reconsiderations during the act of discovery. However, this argument is important because it represents a serious attack on the plausibility of reflexive ideas.

1.5 Background knowledge and experience

As explained before, one of the consequences of trained judgment is the accumulation of experience, a sort of background knowledge that orientates subsequent observation and decisions. For example, a prehistorian accumulates a lot of empirical and practical knowledge about how to observe and excavate this kind of sites. Similarly, Bradley has explained that an effect of having a continuous engagement with some kind of remains is the development of observational powers to notice things formerly unnoticed (Bradley 2003, 153). Hence, the specialists working in a site or a region for a long-term develops a connoisseurship that refines digging and recording procedures (Bradley 1997). However, this form of site-specialization necessarily means being less experienced in other domains.

Likewise, professional diggers accumulate a lot of experience and background knowledge that includes a wider range of sites. Yet, due their focus on fieldwork, there might be some constrains in their background knowledge of finds and artefacts. Moreover, Bradley has explained, that by its destructive nature, the observation of remains is shorter in excavation than in survey techniques. However, this process might be further accelerated by the pace and size of contract excavations and commercial diggers might not have enough time to examine remains and develop their observational powers. In consequence, they tend to observe similar things in every site (Bradley 2003, 155, 2006, 6). If this is a plausible explanation, then these factors should have more weight to account the so-called homogenization of context sheets than simply presuming that standardization is the main cause. Similarly, the process of engagement and familiarization with the site might be affected by the brief participation of commercial diggers in projects which often is determined by short-term contracts, sometimes in a weekly basis (Everill 2009, 13). Under these conditions, then it’s legitimate to ask, what sort of interpretive performance can be expected of fieldworkers that have a very limited period to engage with the site? Altogether, these arguments explain different ways in which the background knowledge affect interpretative work, hence when fieldworkers do no work in a context that facilitates this process, this will constrain the accumulation of experience and the development of connoisseurship.

One of the characteristics of the reflexive methods is having a group of laboratory specialists onsite running preliminary analysis onsite and feedbacking results to excavators. Thereby, fieldworkers will have a wider background knowledge to interpret the site. In general, this mechanism aims to solve some of the problems created by the specialization of fieldwork and the separation of excavation and post-excavation work (Berggren and Hodder 2003, 427). More recently, diggers at Catalhöyük have been provided with tablets to consult previous reports and data onsite, this aims to achieve a similar function of giving immediate access to a database of background information to support field interpretations (Berggren et al. 2015, 442). In this way, the role of background knowledge is not fully unattended in the reflexive method, however if fieldworkers can process all that information at once is unclear.

The role of background knowledge can also be examined in relation to projects. For example, a difference between academic and commercial projects is that scholar projects often work in sites previously explored and therefore there is wider idea of the remains to be excavated (Jones 2013, 180). For instance, when the Catalhöyük Research Project began, there was already a lot of background knowledge about the site due the former investigations of James Mellaart. This previous information was very useful to define research problems and preliminary hypothesis about the function of buildings (Hodder and Farid 2014). However, this doesn't mean that every commercial project starts from zero. For example, a commercial project in London usually has some degree of background knowledge because the region has been frequently investigated and much of these preliminary results have been summarized in research frameworks (English Heritage, 2002). Likewise, some sites might have a lot of background knowledge that enables interpretation before shovelling. A perfect example of this is investigation of the Rose and the Globe given the rich amount of written evidence about Elizabethan playhouses (Bowsher and Miller 2009). However, when background information is more limited or inexistent, evaluation work aims to mitigate these constrains (Carver 2009, 343). Finally, there is another factor that often creates different epistemic scenarios between a research and commercial projects.

Particularly, the recursive relation between field and post excavation stages in long-term research project that facilitates the reconsideration of preliminary ideas (Lucas 2012b, 235).

In sum, background knowledge is an important factor that shapes the epistemic context of investigations and their interpretative results. Initially, background knowledge is defined by the previous experience of fieldworkers, but it is reshaped by the characteristics of their participation in project and the available knowledge of the site. In other words, it's not the same working in a site for ten years than ten months and it's not the same working in Catalhöyük or Sutton Hoo than having to define the basic characteristics of a site that has never been explored. If this is the case, this should provide a better explanation of the descriptive emphasis of some projects and the higher interpretive results of others. And more importantly, a project with more descriptive results shouldn't be necessarily considered bad research.

1.6 The character of remains

The character of deposits is another factor that often has influenced the development of recording systems. For example, SCR was designed to work in deeply stratified deposits (Spence 1993, 25), whereas the feature system was adapted to record landscape sites extended more horizontally than vertically (Carver 2015, 8). These examples also illustrate how the character of remains facilitates and restrains some types of field observations. For example, the extensive nature of landscape sites facilitates the interpretation of spatial aspects (Figure 1.7) whereas this is more difficult in complex stratified deposits (Hammer 2000, 143; Thorpe 2012, 37). The relation between recording systems and the type of deposit also can be observed in graphic records predating the introduction of standardized written formats. For example, the main maps from Star Carr generally show assemblages and distributions of material (Clark 1954, 6). Similarly, the main plans from Yeavering show the distribution of features defining a building, whereas sections usually illustrate individual features or constructive elements (Hope-Taylor 1977, 46). Evidently, these contrast with Wheelerian records in which sections have a primary importance to describe the sequence of sites (Wheeler 1943).

In contrast, one cannot observe any clear influence of British fieldwork traditions in prehistoric or middle east sites in the development of reflexive methods from Catalhöyük. Similarly, reflexive archaeologists usually ignore the character of remains when explaining the success of reflexive methods. For example, Heathrow explorations are usually promoted by its recording procedures largely drawn from the feature system but without giving much care to the extensive nature of the site (Andrews,Barrett and Lewis 2000, 529; Lewis 2006, 17). Only, a few years later, John Barrett briefly commented about the differences of working in a landscape site and a urban deposit (Barrett 2013).

Equally important is the issue of how the preservation and fragmentariness of remains enables or constrains the knowledge of the past (Clark 1954, 74; Lucas 2015, 314). To put a simple example, consider some pottery sherds and depending on its stage of decay, one will have less chances to interpret some aspects. For instance, if the shape is unclear then, there are less chances to interpret the function of vessels. As mentioned before, the survival of remains was crucial in Grahame Clark's interpretative theory, and the opening lines of his Star Carr report also reflect the importance of this aspect.

> *Archaeologists have frequently, and prehistoric archaeologists have as a rule to depend on evidence so vestigial that it hardly does more than delineate the main outlines of the past. Every now and then discoveries are made which, like the tom of Tutankhamen, the frozen tombs of Pazyryk, the Sutton Hoo ship grave, the Royal Tombs of Ur, the painted caves of Altamira or Lascaux, or the moulded and painted shrines of Catalhöyük, illuminate as it were in a flash aspects of the life* [...] *Such finds owe their special character to the completeness with which, due to combination of historical and biophysical circumstances, the material evidence has survived. (Clark, 1954, i).*

Evidently, this is only one side of the argument because technological progress usually improves our interpretative chances by allowing to extract information even of the most fragmentary remains. Nevertheless, fragmentariness can have a more immediate effect in field interpretation and this aspects is also briefly considered by Clark when he compares the excavation of an urban site with standing walls and the exploration of a prehistoric site with marginal indications of decayed structures (Clark 1957, 112). Many years later, post-processual theorists would demand an interpretive theory due the uncertainty produced by the fragmentariness of remains and the fuzziness of layers, but only the second elements was briefly considered (Thomas 1993, 73; Shanks and Hodder 1994, 74–75).

Star Carr and Catalhöyük are good examples to illustrate the epistemic consequences of fragmentariness in excavation. For example, archaeologists of the most recent exploration in the

Mesolithic site have claimed the discovery of a building structure, this interpretation is based in the evidence of a distribution of postholes surrounding a hollow with a dark fill. Besides, the fill was interpreted as a possible hearth after noticing small concentration of burnt flint, although no charcoal or burned organic matter was found (Conneller et al. 2012, 1012). On the contrary, almost any photographic record of Catalhöyük shows very clearly the presence of buildings clearly defined by their standing walls. In other words, the interpretation of Star Carr's structure is less secure for being based in fragmentary evidence (Figure 1.6), whereas the observation of almost any building at Catalhöyük is very secure given the preservation of remains (Figure 1.4).

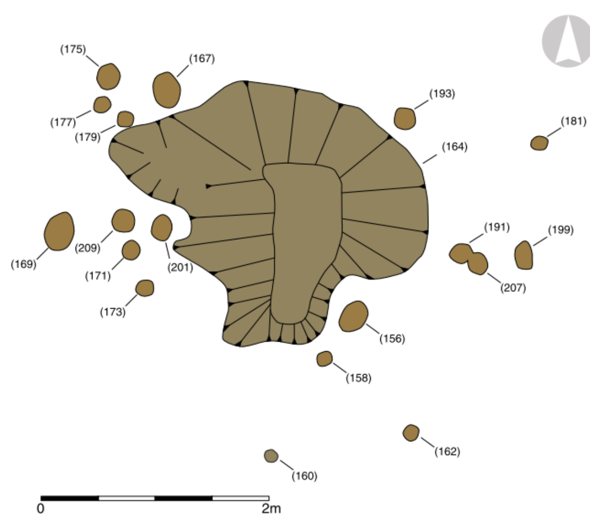


Figure 1.6. Mesolithic structure at Star Carr. Site photography (left), plan drawing (right). (Copyright Star Carr Project CC BY 4.0). https://doi.org/10.22599/book1

Fuzziness in layer boundaries is slightly different because uncertainty is produced by the looseness of sediments, and not by its decay. As Hodder says, layers are not fixed entities of bounded earth and it's uncertain where one layer ends and another begins (Hodder 1999, 85; Shanks and Hodder 1994, 72). However, he uses this argument to question the ontological status of layers (whether these are real or not) without considering its epistemological aspects. Specifically, if fuzziness can diminish the confidence in definitions. For example, imagine that there are two adjacent layers and its boundaries are unclear but there is no doubt about the existence of layers, one is silty and another gravelly. On the contrary, consider an alternative scenario with an accumulation of silty deposit, but one portion has a small amount of gravelly inclusions and therefore is not clear whether there is one or two layers. Finally, one must take account that this sort of fuzziness is contingent because there will be many deposits with very clear boundaries too, for example a floor made of stone slabs.

These philosophical games are important due their implications for recording activities. Reflexive archaeologists say that one must document the process of discovery, otherwise the credibility of data is compromised (Hodder 2005, 651; Mickel 2015, 303). However, the relevancy of this type of testimony is not very clear, especially if one considers that photographs can provide a sound testimony of the existence of buildings at Catalhöyük. On the contrary, an account of the interpretation of Star Carr's structure is most necessary because a graphic testimony is not conclusive. Similarly, in the case of layers, it's pertinent to think about the characteristics that documentation should have when fuzziness affects the credibility of layer definition and when it does not. In short, if the character of remains shapes the interpretative process and affects the credibility of data, ideally, these aspects should be considered in the process of documentation. Once this has been stablished, it's legitimate to consider what are the interpretative and recording differences when working in a well-preserved site like Catalhöyük, a landscape site with no deep stratigraphy and a multiperiod site in London with more fragmentary remains.

1.7 The role of excavation strategy

Excavation strategy is another element that shapes the interpretative process of fieldworkers as it enables to see more easily some things instead of others. For example, Flannery explained how extensive explorations facilitates the interpretation of spatial relations among features and assemblages (Flannery 1976, 50). Similarly, Barker recommended open areas for the detection of timber structures that otherwise would be less evident by test pitting and trenching (Barker 1969, 233). On the contrary, excavation pits and trenches are strategies that facilitate the interpretation of stratigraphy. This doesn't mean that stratigraphy is invisible in plan, but this requires more attention to be perceived.

In general, strategies are defined considering the character of remains and the research agendas of projects. For instance, Barker's most emblematic extensive excavations took place in relatively shallow deposits and despite his great support for open areas, he accepted that trenching would be useful for observing constructive phases in stone buildings (Barker 1969, 233). Other paradigmatic examples are the extensive exploration of landscape sites for interpreting marginal wood structures at Yeavering (Hope-Taylor 1977, 47) and the deep trenches at Jerusalem to follow constructive sequences (Kenyon 1974, 55). Likewise, strategy might be also applied at lower scale when dissecting a feature. For instance, burials are explored extensively to observe a horizontal patterning of artefacts whereas postholes will be sectioned to observe vertical relations of remains. In general, the incidence of strategy is more evident when one compares different strategic approaches. However, this aspect will not be further considered in this investigation because all the selected case studies implemented the same strategy: extensive excavations. This selection was intentional to facilitate comparability.

Finally, there is one aspect that deserves a short note. One of the characteristics of commercial projects is having less control over strategy since excavation areas are largely defined by development (Jennings 2011, 285–286). This aspect will not be further revised either, but a recurrent problem in the commercial sector is the limited interpretative potential of small-scale excavation like test-pits and tunnels. However, this doesn't mean that these data are useless. Patrick Ottaway analysed the primary records of watching briefs, test-pits and open areas in York, and then he achieved a general understanding of Roman York (Ottaway 2011).

1.8 Collective empiricism: The team and the record

Many theoretical discussions about excavation and recording are grounded in a distinction between objectivity and subjectivity. Followers of post-processual ideas often criticise the mechanic depiction of excavation that obscures its subjective components (Chadwick 2003, 99; Adams and Brooke 1995, 96). In some extend these comments are pertinent, because methodological discourses usually stress mechanical aspects, but this doesn't mean that fieldworkers don't acknowledge the importance of subjective aspects like skill and experience. Still, some interpretative theorists strongly believe that there is a campaign in British archaeology to promote objectivity at the expense of subjective elements (Hodder 1997, 692). On the other hand, the post-processual stance has its own deficiencies as subjectivity is not very well defined and they fail to distinguish between an opinion grounded in skill and experience and an opinion based in limited skills. In consequence, they create a mutually exclusive relation between subjectivity and objectivity. However, as explained before, mechanical objectivity and subjective trained judgment generally work together. Then, the irony of British methods is that standardized procedures work better when refined by subjective experience whereas reflexive practices must be systematized to become operable.

Another characteristic of recent theoretical work is the definition of an interpretative theory largely based in individual agents of knowledge, most likely due a strong emphasis in subjectivity and embodied cognition (Hodder 1997; Chadwick 2003; Bender et al. 2007; Shanks and McGuire 1996; Thomas 2004). Generally, this stance takes examples of 'seeing different' to demonstrate the subjective components of perception and interpretation (see 1.2). However, this view fails to acknowledge that cases of seeing similar are more frequent because people share much knowledge (Hanson 1958, 18). Besides, this is closely related with the standardized language of objective representations that aims to produce disembodied representations that can be understood by an epistemic community (Edgeworth 2003, 99). In other words, one of the limitations with recent interpretative theory is failing to consider that observers are not isolated, they are participants of social groups (Radder 2006, 82).

Another reason to revise interpretive theory is the nature of fieldwork which primarily is a collective enterprise or teamwork (Figure 1.1). In the first place, fieldwork is collective because it requires a large workforce due the division of labour and the scale of excavation work (Olsen et al. 2012, 68; Carver 2016, xiv). However, the interpretative implications of collective work have not been considered very often. Once more, Edgeworth is an outstanding reference, because although most of his work describes the embodied dimension of interpretation. He also stresses a collective component in interpretative practice that he defines as 'social transactions' (Edgeworth 2003, 41). For instance, Edgeworth describes how two or more diggers coordinate their movements when working together in the definition of a feature. Specifically, he classifies this type of collective interpretation as joint actions and he remarks that many of these exercises can be carried out in silence due the existence of a common background knowledge about observable entities and technical procedures. However, there is always a constant sharing of empirical experience as fieldworkers continuously take account of the discoveries made by their partners (Figure 1.7). Evidently, there are situations that might require dialogue, for example to direct attention to certain aspects or discussing conflicting evidence (Edgeworth 2003, 41–42). Additionally, Edgeworth considers another form of collaborative exercise which is mainly mediated by language. For example, he mentions that informal trench visits create opportunities for observing the latest discoveries, sharing previous experiences and discussing the meaning of evidence (Edgeworth 2003, 44). Evidently, this mechanism has been refined in reflexive methods in the form of 'priority-tours' to regulate the exchange of information and discussion of evidence (Figure 1.8). Hence, this interactivity aims to promote a collaborative work (Hodder 2005, 653; Farid 2015, 67).

In sum, joint action and interactivity characterize two epistemic scenarios of collaborative interpretation in which the individual perceptions and previous experiences of agents become public or part of an epistemic community. This is possible because there is a set of common assumptions that defines what things are worth seeing and a standardized language that stablished how to describe and communicate findings. These elements indicate that interpretive fieldwork is performed within a social context in which a group of agents produces knowledge working together. Overall, this constitutes the mains interest of what has been defined as social epistemology (Goldman 2010, 14). For this reason, an interpretive theory strongly grounded in subjective experience and embodied cognition is insufficient.

Figure 1.7. A group of diggers defining a feature. Heslington East, University of York (Copyright York Archaeological Trust, CC BY 4.0). https://doi.org/10.5284/1019860

Figure 1.8. Priority tours for discussion site interpretations at Building 49, Catalhöyük. (Copyright. Catalhöyük Research Project. CC BY-NC-SA 2.0)

In the past decades, the issue of the completeness and selectiveness of site records has been a crucial topic in theoretical discussion (Carver 1989). Much of that debate emerged in the context of rescue archaeology, particularly to question the policy of preservation by record (Lucas 2012b, 62). Critical responses often suggested that field data should be recorded within a research framework (Binford 1964; Reynolds and Barber 1984). However, this idea was untenable for some archaeologists because it implies producing an incomplete record that ignores relevant data for other researches (Flannery 1982, 275; Barker 1982, 37). Evidently, asking for completeness should be understood relative to a set of contingent standards for an epistemic community. However, another important purpose of this view was achieving an empirical goal of sharing data within an epistemic community. In contrast, supporters of selectivity generally believe that if records only make sense within a research framework, then it's almost impossible to reuse an archive for an alternative agenda. However, this argument is contradicted by various examples of recent investigations using antiquarian observations, for example (Barrett,Bradley and Hall 1991). Hence, despite antiquaries didn't record the same things that modern prehistorians would, the last can use antiquarian observations because there is a common cultural background about things that deserve to be observed, for instance monuments, artefacts and landscapes.

More recently, reflexivity has questioned the completeness of standardized archives by suggesting the lack of crucial contextual information. Hence, reflexivity aims to create a more complete archive that fills those gaps. This is a bit paradoxical if one considers post-processual critics toward completeness ideals (Tilley 1989). However, reflexivity moves in the same direction by enlarging the amount of recording almost in an attempt to capture the whole experience of digging a site (Chadwick 2003, 103; Lucas 2012b, 72). Nevertheless, one must acknowledge that a primary goal of reflexivity is improving the analysis of records, which in theory is very problematic with a set of standardized documents.

The archives produced over recent decades have increasingly become problematic since they are not sufficiently contextualized within a reasoning process. Many are extremely difficult to use. Diaries are often not present and unless one can talk to the individuals who were involved in the original excavation, it is often difficult to make sense of large amounts of highly formalized data. The provision of a reflexive context must become one of the main aims of methodology so that it is possible to make sense of archaeological fieldwork after the event (Hodder 1999, 31).

Again, there is contradictory evidence against reflexive claims, specifically there are various examples of backlogged standardized archives analysed and published many years later (Bowsher and Miller 2009; Evans et al. 2016). Besides, these examples are relevant to indicate theoretical limitations in reflexive discourse. For example, one of the basic lessons of historiography is that documents only speak when they are properly interrogated (Bloch 1954, 53) and this an analytical ability is another interpretative skill (Ginzburg 1999, 20). Hence, this indicates more clearly another possible theoretical inconsistency in reflexive discourse because although they observe similarities in the process of interpreting archaeological remains and documents (Hodder 2000b), they deny the possibility to examine context sheets. Hence, it’s possible that standardization and the lack of information are not the main problem of formalized records. Instead, it’s possible that the value of standardized data is unclear due the lack of an analytical strategy to make them speak.

One characteristic of documental analysis is that testimonies must be confronted with other documents to interpret their meaning and check their accuracy (Ginzburg 1999, 12–13; Bloch 1954, 54). Thereby, documents are always examined in series. In some extent, reflexivity coincides with this idea as they consider that context sheets must be confronted with diaries. But this process of cross-examination is neglected for the sequence of context sheets with no good reason. This is a highly problematic because context sheets are connected in multiple ways, for example by stratigraphic and feature relations. In other words, context sheets form a contextual network of primary sources that enables analysis by cross-examination. Not to mention that the sequence of context sheets is inserted in a wider network that includes graphic testimonies. In sum, there are two additional theoretical problems in reflexive discourse. First, they fail to notice the contextual network of context sheets and secondly, they miss the relevancy of documental analysis. Moreover, the analysis of any form of record is possible because these sources are recorded with a public framework defined by disciplinary standards. Even, if this is not completely true because there might be variations among different traditions of archaeology, the task is possible by studying the language and practices regulating observations and representations. Evidently, this is not a simple task if primary data are fragmented into multiple documents over a sequence, but the task is possible by cross-examination. These ideas constitute the core of a method for documental analysis of site records to be described in the next chapter.

Chapter 2 Methodology

The main concern of this investigation is what has been defined as the problem of quality in recording methods for archaeological excavation. Broadly speaking, the problem consists in investigating whether there are improvements in the variability and credibility of excavation data when projects implement more complex recording methods. However, as mentioned before, there is an alternative way to explain quality differences in excavation records by considering the incidence of the epistemic context in field investigations. This study selected three projects that implemented alternative recording strategies. Specifically, The Catalhöyük Research Project (CRP) which represents the reflexive method, the excavations at MRG95 a typical commercial project in London recorded with single context recording and the Tarbat Discovery Programme (TDP) which followed the guidelines of the feature system.

These projects were selected due their recording strategies and context in which they took place, namely the academic and commercial sector. This variable was important as it had implications in the type of workforce, the duration of projects and the background knowledge of sites which altogether defined the epistemic context of projects. Section 2.1 expands on the selection of cases studies. Generally, debates about the virtues and problems of recording systems have been largely unfold without paying much attention to empirical evidence. Hence, case selection was a crucial step for defining documental samples to be studied. Then by looking into the content of records, this study aims to give a more reliable answer to such matters. Section 2.2 provides a general account of the type and size of samples.

The analytical work is divided in two steps. First, there is an internal analysis of case studies followed by a comparative step. Section 2.3 describes the qualitative and quantitative procedures for the internal analysis of samples to describe the variability, credibility and relevance of data, additionally it describes the process to examine the incidence of circumstantial factors like the skill of fieldworkers, the background knowledge of projects and the character of remains. In short, the aim of the internal analysis is getting a detailed picture of the interpretive process of case studies. Section 2.4 describes the quantitative and qualitative aspects for the comparative phase. However, the internal and comparative stages are based in an elementary analytical strategy defined as content analysis or coding. Content analysis is a procedure to classify the content of excavation records which was crucial to describe the variability of data. Coding was guided by a classificatory system organized in classes and types. Section 2.5 explains the process of coding and Section 2.6 explains the logical and conceptual background of the classificatory framework.

2.1 Case selection

Case study is a research strategy to analyse situations and events in relation to their context. This strategy is used in social sciences and management studies to investigate different ranges of social phenomena and the performance of companies and projects, respectively (Yin 2014, 31). However, this inductive strategy has a longer tradition in historiography to study individuals and events in their surrounding but always aiming to reach general conclusions (Ginzburg 1999, 12; Levi 2001, 98). Equally important is noticing that the context of a case study can be defined by social elements as class and gender relations or environmental factors in which a person or a group reacts. Case studies were selected for being representative of different epistemic circumstances and recording methods, these differences were crucial for the research questions of this investigation.

The first case study is the Catalhöyük Research Project (CRP), which was one of the main experiments on reflexive methods. It was a long-term research project in the famous Neolithic site in Turkey. The field phase began in 1993 and finished in 2017 with annual seasons every summer. The project included two main excavation areas (Figure 3.1) and the site included roman and byzantine remains, but the main occupation dates from the Neolithic. The excavation team included a mix of professional diggers and students (Farid 2015). Besides, when the CRP started, the site was already famous due the former exploration of James Mellart that had already defined the architecture of the site (Mellaart 1967). Then, although the CRP didn't take place in Britain, the methodological and operative core was deeply British.

The second case study are the exploration at MRG95, this was a typical commercial project in the city of London executed by MOLA (then MoLAS). The excavation phase lasted for a few months between February and May 1999 and was executed by a team of professional diggers. The project only included one extensive excavation area and the site covered a multiperiod deposit ranging from roman to post-medieval deposits located in the Walbrok valley (Figure

4.4). This area was known for being the main industrial area of Roman London (Seeley and Drummond-Murray 2005, 5–6).

Finally, the third case study was the Tarbat Discovery Programme (TDP), another long-term scholar project that investigated the site of Portmahomack, an early medieval deposit located in the Tarbat Peninsula of Scotland. The project included four main excavation areas or sectors (Figure 5.3). Fieldwork run between 1994 and 2007 and included the participation of professionals, students and volunteers. The site of Portmahomack is a deposit that extends across the landscape without a deep stratigraphy (Figure 5.4). It includes the remains of a pictish monastery and although pictish monuments had been previously studied, no pictish settlement had excavated before Tarbat's explorations (Carver 2016, xiv–xv).

The definition of case studies had to consider the size of projects as the CRP and TDP included more exploration areas than MRG95. Hence, units of analysis were constrained to specific extensive explorations to facilitate collection and analysis work. Specifically, these include the exploration of Building 49 (B49) for Catalhöyük (Figure 3.3), Sector 2 for Tarbat (Figure 5.5) and the only excavation area for MRG95 (Figure 4.5). MRG95 was the earliest excavation and took place in 1999 during the PPG16 era. Tarbat's exploration took place between 1996-2007, and B49 was excavated between 2004-2008. This temporal proximity was intentional to compare relatively contemporary projects. Finally, these cases were selected for having published monographs, this was an important factor for two reasons. First, it was considered a sign of well-funded projects that allowed to minimize the incidence of economic factors. And secondly, reports provided background information about projects and sites, which was useful during the process of case selection and sampling. Evidently, this process of selection limits some possible contrasts, for example the implementation of reflexive methods in commercial scenarios (Lewis et al, 2010). However, the selected projects work well for the main goals of this investigation.

Case study	Recording system	Type of project	Workforce	Character of the site and background knowledge
B49	Reflexive method	Catalhöyük Research Project. Long term research project	Mixed team with professional diggers, graduate students and trainees aided by laboratory specialists onsite	Catalhöyük. Urban Neolithic site with a high preservation of remains, famously known for the repetitive pattern of buildings across the site
MRG95	Single context recording	MRG95. Short-term commercial project	Team of professional diggers	MRG95. Multiperiod site with different levels of preservation located in the industrial area of Roman London
Tarbat	Feature system	Tarbat Discovery Programme. Long term research project	Mixed team of professional diggers and trainees	Tarbat. Early medieval site located in the vicinities of the Monastery of St. Colman. The deposit covers an industrial area with different levels of preservation extended over the landscape.

Table 2.1. The table resumes the main criteria for case selection.

2.2 Sources

As mentioned before, recent methodological debate in Britain is largely based in professional opinion, except for a recent study that actually has examined reflexive diaries (Mickel 2015). For this reason, it was important to examine a sample of primary records produced under alternative methods. The use of documental evidence for investigating philosophical questions is inspired in the notion of a naturalized epistemology or the idea that philosophical research should be supported by empirical evidence and not simply by argument and conceptual analysis (Quine 1969, 69). Therefore, by studying primary records one should be in a better position to evaluate the merits of principles. As explained before, written formats constitute the focus of this study, particularly excavation diaries and sheets. However, context and feature sheets also include graphic data like stratigraphic diagrams and sketches, hence these textual and graphic data were analysed altogether.

Also, as explained before, one of the differences between the reflexive method, SCR and the feature system is the number of sequences of written formats. Although, the extension of sequences varies. For instance, sequences of contexts and feature sheets reach thousands and hundreds respectively. Whereas excavation diaries only include a few tens of entries. Still, these represent a lot of sources to be examined, then they had to be sampled to facilitate the analysis. Around 25% of the sequences of MRG95 and Tarbat were collected, these samples aimed to be representative of the different stages of the project and its participants. Hence, the size and distribution of samples enabled to analyse the general results of the project, the individual performance of fieldworkers and the flow of interpretative work through time.

In the case of B49, only the full sequence of diaries was examined. This decision was made considering the existence of a recent study that already compares the content of sheets and diaries from the reflexive method (Mickel 2015). This study stablishes similarities and differences in the content of both formats, but in general it stablishes a stronger descriptive character in context sheets and a stronger reflexive nature in diaries. In general, these results seem plausible, then I was not interested in replying that comparation. The most controversial aspect is that investigation is the interpretation of results, specifically the explanation of diaries' content enables to stablish the credibility of site interpretations (this aspect will be considered more extensively in section 3.6). This constitutes an important reason for prioritizing the study of reflexive diaries from Catalhöyük. Another questionable aspect is Mickel's argument is that after comparing the content of context sheets and excavation diaries from Catalhöyük, she stablishes a general contrast between SCR and the reflexive method. This interpretation is controversial because she mainly compares the content of two formats (context sheets and diaries) within a recording strategy (the reflexive method). A fairer comparation with SCR would require considering the results from an archive that only included a sequence of context sheets. As explained before, the difference between SCR and the reflexive method rest in the number of written formats that each strategy includes. For this reason, my investigation chose to compare archives from different projects rather than simply compare context and diaries from the same project.

2.3 Internal and comparative analysis of case studies

The analysis of records is organized in two steps. First an internal analysis of case studies and secondly a comparative stage. The internal analysis was grounded with a mix of qualitative and quantitative procedures for the analysis of records. The main goal of the internal analysis was to produce general descriptions of projects based in the variability, reliability and the relevance of data. This information facilitated the analysis of the interpretative process at different scales: the project, the individual performance of fieldworkers and across different stages of the project. This sum of qualitative and quantitative observations enabled to get a detailed picture of the incidence of epistemic factors, especially the skill of fieldworkers, the character of remains and the role of background knowledge. Overall, the internal analysis aims to be an exhaustive study of projects, grounded in an intensive study of documentation (Levi 2001, 95; Ginzburg 2012b, 207).

Coding

Content analysis was the most basic form of qualitative analysis from which further forms of analysis derive. Content analysis or coding consist in classifying the informative content of excavation records into various classes and types of information (Silverman 2011, 164). Coding was guided by a conceptual framework organized into three classes: interpretative, reflexive and non-interpretative information. And each class was integrated by multiple types of information which represents the most basic form of data (Table 2.2).

Quantitative observations: variability and frequency

After classifying the content of records, the frequency of the distinct types of claims was counted to observe the general distribution of claims in every project. This process allowed to observe the proportion of layer descriptions, reflexive data and so forth in every project. Afterwards, the performance of fieldworkers was by examining the distribution of claims by fieldworker. This data allowed to observe if some diggers had produced more interpretative and reflexive testimonies than others. These results were interpreted as differences in the skill and participation of fieldworkers. Finally, there was a longitudinal analysis to examine the distribution of claims across different stages of projects. Three arbitrary stages of were defined in case studies, although the length of stages varied in relation to the duration of projects. The longitudinal analysis aimed to study similarities and differences in the type of data recorded at different stages of a project. Arguably, this form of analysis would allow to observe whether changes in the background knowledge of a site would enable the interpretation of high-order claims.

Qualitative analysis is based in elementary procedures of descriptive statics, basically the frequency and percentage of attributes in samples. No more complex statistical analysis was included, for example inferential static, because one of the basic goals was observing quantitative similarities and differences and study them in relation to qualitative aspects defined by the context of case studies, not to define general statistical trends in which conditions are unimportant. Moreover, the size of some samples (B49 for example) is not too extensive to be subject to more complex statistics. There are some kinds of statistical tests that could be applied for some quantitative results, for example to assess dependence relations between variables. However, this outreaches the scope of my study. Hence, the focus will be in defining significant attributes and relations that could be subject to more complex statistical analysis in the future.

Qualitative observations: relevance and validity of data

Additionally, there were additional forms of qualitative analysis to examine the relevance and validity of specific types of information in records. Reflexive and non-interpretative data were examined to assess whether these types of claims contributed the credibility of interpretative data. Hence, reflexive and non-interpretative information was classified in different categories according to their relevance. This information allowed to observe more clearly how much of that information was important and how much was unnecessary.

The validity of testimonies was exclusively addressed for interpretive claims in excavation sheets. Hence, this was exclusively applied to MRG95 and Tarbat samples. Specifically, this analysis aimed to evaluate the credibility of interpretations in recording methods that do not include diaries. According to reflexive archaeology, this task is impossible (Hodder 1999; Mickel 2015). In broad terms, assessing the credibility of interpretative data requires checking its empirical support. For example, if a sheet interpreted a layer as a dump, then one had to examine whether that record described evidence supporting that claim, for example a conjecture that explains how descriptive and interpretive data link. In many cases, this only required checking the content of a sheet, but in some other cases it required confronting the content of two or more sheets. This cross-checking of different testimonies is a typical historiographic methods to evaluate the credibility of documents (Ginzburg 1999, 12–13). Evidently, this qualitative examination allowed to differentiate between reliable and unreliable sources, however, this is very basic method to evaluate the credibility of interpretive data because the strength of interpretation was not considered, for example whether a claim was supported by sound or weak evidence. In the end, these qualitative observations enabled to produced additional quantitative data, specifically the proportion of relevant reflexive and non-interpretative data, and the proportion of reliable sources.

Comparative analysis

The comparation of project was the following analytical step to observe similarities and differences in the interpretative process of sites and the quality of testimonies. This contrast is based in a comparation of the general distribution of claims of case studies, this contrast was useful to identify whether some projects had produced more interpretation and reflexivity than others. Nevertheless, these results were interpreted considering the epistemic context of projects. The comparative stage was complemented with a contrast of the relevancy of reflexive and non-interpretative information of case studies. Finally, there was a comparation of the reliability of excavation records for MRG95 and TRS2. This contrast enabled to identify the areas of field investigation in which fieldworkers were more and less successful generally. Overall, the internal and comparative judgment allowed to observe more clearly the incidence of methodological and epistemic circumstances in the quality of records.

2.4 Basics of coding

As mentioned before, coding aims to classify the information of records into different classes and types (Silverman 2011, 164). This process was guided by a conceptual framework that helped to identify the variability of information in documents (Table 2.2). In practice, coding implies fragmenting the content of a diary entry or a context/feature sheet into smaller pieces of information. The fragment below is a typical example of coding for a diary entry from B49, Catalhöyük. Initially, the entry describes the process of layer definition of a fill and some related operative procedures. It continues with a description of the discovery of some structures followed by additional operative information related to the removal of the fill. Finally, there is an reflexive claims describing some doubts related to the discovery of structures.

> *We have been continuing work on unit 7913, the giant fill layer.* (**Context definition**) *We have fast-tracked this unit and are steadily going through it, pedistaling anything that looks remotely interesting* (**Operative decision**). *We started to follow the lines of bins (?) from unit 7916 out of the fill to find the outline of the structure(s). We think that there are some bins in the south-west corner. Thus far, we have just come down upon the plastered outlines of the bins* (**Structural element**). *Today, we (DGN & SHL) began to dig westwards in an attempt to remove the fill bulk. We stumbled upon several plaster lumps that later Ian and Shanina told us to take out, based on previous experience of these types of house fills* (**Operative decision**). *By the end of the day, I think we have at least two- possibly three- bins outlined or we think we know where they go. I exposed the plaster on the outside of one bin abutting the east and north wall. I think there are two other bins in the area exposed by unit 7916* (**Uncertainty**). (SL/02/04)

Coding context and feature sheets was slightly different because information is recorded in different sections. Then, in some extent, fragmentation of data has been already implemented. Overall, this facilitated coding sheets. However, the interpretative section of sheets usually contained different types of claims, that had to be coded into smaller bits of information (Figure 4.1). In theory, classificatory categories must be sufficiently clear to be mutually exclusive and avoid overlying and double coding. However, this was unavoidable sometimes due the close textual and conceptual relation between reflexive and interpretative claims. Specifically, reflexive claims (uncertainty, reconsiderations and possibilities) are always about some interpretative aspect (see Table 2.1). Occasionally, records included an interpretative claim and later if required, a reflexive statement was added as in the example above, but this was not always the case, and sometimes interpretative and reflexive data conflated. Finally, it should be said that coding was selective. Specifically, referential data (dates, names and site codes) was ignored from records for being irrelevant for analytical purposes. Coding was implemented with the aid of a software (NVIVO), which allowed to create a database with the coding process of each document. This was the basis form secondary databases that include quantitative data (see accompanying materials).

A conceptual background for interpreting the meaning of records

The codification of records is an analytical exercise derived from interpreting the meaning of information. Hence, the most immediate need is having a strategy for the accurate interpretation of sources. To achieve this task a strategy in two steps was developed. The first step consisted in having a conceptual background to guide the interpretation of sources; and secondly, the development of a contextual strategy to interpret the meaning of data widespread over a document and organized in sequences. As the process of coding started, it was evident that a lot disciplinary language regulates the meaning of data. This was most evident in layer and stratigraphic claims where terminology is highly controlled. Then, the first step to make reliable classification was to define a conceptual framework. In some extent, many of these notions were already familiar and are related to the theoretical framework of stratigraphic excavation. Still, conceptual revision was useful to identify notions with multiple meanings like the assemblage that sometimes is described as a group of finds within a context and sometimes refers to a patterned arrangement of artefacts (Joyce and Pollard 2010). Likewise, it was necessary to study the codes of representation, especially for the interpretation of graphic data like diagrams and sketches in excavation sheets. Again, many of these codes were already familiar, but other requires further study to be fully aware of the meanings of the different types of lines to represent uncertain boundaries and truncated contexts for example (Museum of London 1994).

Finally, another challenge was studying the use of less standardized information. This was more evident with formation processes and spatial data. For example, some records describe formation processes in terms of primary and secondary deposits, but others described them in terms of construction, destruction and dumping events. In consequence, having a conceptual background helped to justify the classification of this information into a single category. Overall, conceptual revision showed that a lot of the information is regulated by shared notions, mainly because as Wittgenstein explained, language is a communal practice (Brand 1979, 110). Then, when the meaning of something was unclear, the initial strategy consisted studying the local terminology to learn the forms of expression of a general concept.

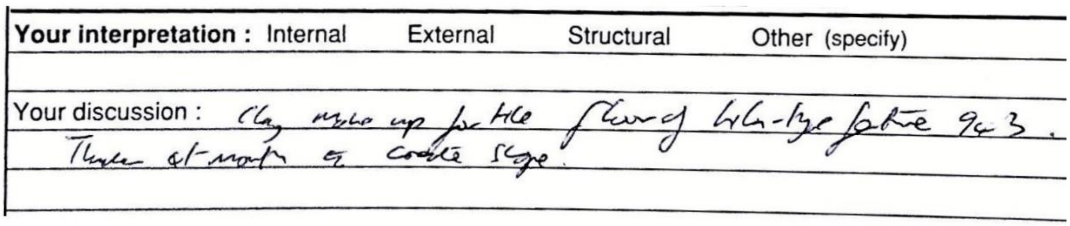
A contextual strategy for interpreting the meaning of records

Sometimes an additional step was necessary, and this required reading a piece of information in relation to another to interpret its meaning. Sketches included short explicative notes occasionally, but when they didn't, linking textual and graphic information was useful to interpret the meaning of drawings, for example if a sketch represented a structural remain (Figure 4.2). Additionally, many sheets included updates, which often produced small pieces of disconnected text or graphics. Then, after linking those separated bits the meaning of text made more sense and frequently this allowed to observe fluidity of interpretations (Figure 4.18). Another interpretative strategy consisted in reading sources as part of a conditional sequence, then by following the contextual links between one sheet/diary entry and another this enabled to understand the meaning of data more clearly. Moreover, this strategy also allowed to observe reconsiderations among testimonies, either textually or graphically recorded (Example 3.5 and Example 4.14). These interpretative strategies are highly consistent with the principles of contextual archaeology (Hodder 1986, 119). However, I have drawn more closely from the notions of documental analysis as portrayed by historians, that insist in the need to analysed series of related documents, either to interpret its meaning or to check the validity of testimony (Ginzburg 2012b, 202; Bloch 1954, 92).

Atypical meaning problems

One of the characteristics of excavation sheets is being handwritten documents and this produced some minor difficulties to interpret the meaning of words. This included the use of abbreviations, although some of these were standardized such: '*occ*', '*med*' and '*freq*' for describing the quantity of inclusions. In other cases, fieldworkers developed their own codes such as: "*cut into b/earth surface 219*" (HK-256). However, the immediate context generally provided enough information to interpret its meaning. More problematic was unclear handwriting, this was more recurrent in the records from one person at MRG95. In this case, palaeography was useful, but sometimes the immediate context of an illegible word or statement was useful to interpret its meaning, including the information of related sheets recorded by other participants. Still, there were a few sheets with some bits of information that couldn't be deciphered. For example, the first line in Figure 2.1 was interpreted as: "clay makeup for the floor of kiln-type feature 943"

(JDM-1022). However, it was impossible to reach a conclusive interpretation for the second line. Overall, the number of sheets with undecipherable bits of information was less than five, then being statistically insignificant.


Your interpretation : Internal External Structural Other (specify)

Your discussion : [illegible]

Figure 2.1. An illegible testimony of a context sheet

In the case of diaries, referencing problems were more frequent. For example, an entry describing the excavation of a horncore define it as ‘the possibly attached horncore’, but other entries referred to the same element refer to it as horn-core U.7920. In this case, the context of entries suggested that both testimonies referred to the same thing, and this was confirmed with one testimony that included both references (Example 3.5).

Finally, there were some sheets with ambiguous terminology particularly, for example when using the notion of association to describe different types of interpretation. One diary entry uses this term to describe an activity area and later its functional interpretation: “*Basin F4010 was located on a platform in the NW corner and probably associated with an area of charcoal and dirty floors. This area appears to represent an activity area, possibly associated with food preparation*” (DE/46/08). Likewise, a context sheet also used the same term to describe a group of features: “*stakehole associated with other stakeholes*” (HK-254), but in this case the meaning was confirmed after revising an additional sketch (Figure 4.28). Finally, another sheet mentioned: “*ashy deposit probably unrelated to 700. More likely to be a dump from cleaning kiln or furnace*” (MW-710). Initially, the meaning of being unrelated was unclear, however an accompanying sketch illustrated that layer 710 was possibly inside cut 700, then the type of relation being described became clear. Altogether, these examples represent different instances when contextual information had to be used for interpreting the meaning of ambiguous statements. Overall, there were less than five sheets in MRG95 and Tarbat’s samples with ambiguous claims that couldn’t be deciphered, therefore being statistically insignificant.

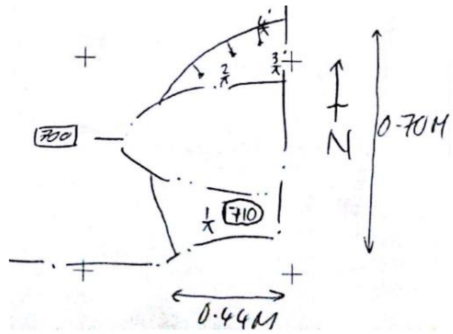


Figure 2.2. Sketch from a context sheet (MW-710).

2.5 A classificatory framework for sources

A classificatory framework was defined to guide the process of coding sources. It is organized in classes and types. It includes three classes with multiple types (Table 2.2). The main classes are non-interpretative, reflexive and interpretative claims. Interpretative claims include information related to entities such as layers, features and structures, but it also includes stratigraphic information and formation processes among others. Reflexive statements describe doubts, ambiguities and reconsiderations related to interpretative data. For example, if there is some sort of uncertainty when defining a layer or interpreting the function of a structure. Finally, non-interpretative statements include different forms of background information. For example, observational conditions and operative decisions related to excavation work. Then, although treated as independent classes, there is some intersection of reflexive and non-interpretative claims with interpretative statements.

Furthermore, the class of interpretive claims is organized in three hierarchical levels: low, middle and high-level claims. This arrangement is based in the conceptualization of fieldwork as interpretative work, hence activities of data production based in observation like layer definition are considered low-level claims whereas the interpretation of unobservable phenomena like functional aspects are represent high-order claims (Hodder 1999, 81; Thorpe 2012, 36; Hammer 2000). This hierarchical order is based in the evidential relation between levels of interpretation, for example the interpretation of a feature depends on the interpretation of individual layers. This hierarchy has been sketched by Martin Carver when he explains that a group of contexts define a feature, and group of features define a structure (Carver 2009, 19–20). These ideas backup the definition of a basic structure of three levels of interpretation:

- Low level claims. Include statements related to the definition of basic components such as contexts and finds.
- Middle level claims. Include statements related to the definition of secondary entities like assemblages, features, structures, groups of features, and building. It also includes claims about different types of depositional processes like stratigraphy, formation processes and constructive phases.
- High-level claims. Include interpretations of social and cultural aspects of the past. For instance, interpreting the function of a building or the cultural origin of a feature.

This arrangement reveals that lower-level claims are closer to direct knowledge or empirical experience. High-level claims are inserted in the domain of inferential or indirect knowledge. Whereas middle level claims oscillate between direct and inferential knowledge. This clarification is important because it highlights divergences with Carver's ideas as he mentions that "a feature doesn't exist physically; it is a set of contexts (which do) […] The structure exists at still higher level of interpretation and still less material, more of an idea, than its features" (Carver 2009, 140). This is problematic because there are not good reasons to assume that a burial might be less real than its fill. In my view, a context is as real as a burial pit. Evidently, there might be cases in which the interpretation of a building based in alignments of potholes might not be very secure if evidence is marginal, but in that case, uncertainty will rise from the quality of evidence, not from the conceptual status of entities. Likewise, it should be mentioned that my hierarchical system doesn't not coincides with Hawke's ladder of inference which aimed to emphasise the level of security of interpretative claims, then in his view some types of claims are more secure than others (Hawkes 1954). Hence, this hierarchical arrangement doesn't presume that lower-level claims are more secure than high-order claims, because empirical experiences are as fallible as inferences.

Most of this classificatory system derives from the revision of concepts from British archaeology. Especially, the notions of stratigraphic excavation which include concepts about layers and stratigraphy hic relations (Harris 1989). Other notions like the feature and the group of feature have been imported from the conceptual tools of the feature system (Carver 2009, 19–20). Still, some conceptual references from American and French archaeology have been incorporated when no better definition among British literature was available. Some types of claims are based in preliminary notions enriched by documental analysis, particularly for reflexive and non-interpretative claims. Finally, some types of claims emerged inductively when examining sources. It's almost certain that there will be some limitations in the applicability of this framework beyond the domain of excavation. For instance, this system proposes that space and building claims are higher order than context definitions. Yet, this might not be true from the point of view of survey techniques where spaces and buildings might be the most basic forms of claims whereas context and stratigraphy might be less accessible (Bradley 2003; Jones 2000). Next, a detailed description of the conceptual framework and its types will be presented. However, a synthetic version is included at the end of this chapter (Table 2.2).

Non-interpretative information

The class of non-interpretive information includes three types of claims: Observational circumstances, decisions and ethnography claims. Observational circumstances include descriptions of observational conditions related to the interpretative process, for example whether excavation had taken place under artificial light or if a layer had been observed in section but not in plan. Meanwhile, decision claims include different forms of operative decisions, for example if a layer had been sliced into arbitrary levels. Finally, ethnographic claims describe activities related to the project but not to interpretative work such as having visitors onsite. Altogether, these claims provide different forms of contextual background for reflexive and interpretative data. In some extent, these information corresponds to what Hodder and others have described as providing contextual information about excavation work (Hodder 2000a, 9; Mickel 2015, 300). However, this framework follows more closely the concepts related to notion of material realization in observational acts (Radder 2006, 31).

Reflexive information includes five types of claims: uncertainty, fluidity, possibility, interactivity and methodology. In broad terms, reflexive claims describe doubts, ambiguities and reconsiderations related to the interpretative process. These notions are imported from the hermeneutic model and the theory of the act of discovery (Hodder 1999, 36; Edgeworth 2003). However, reflexive types are nourished by discussions regarding the doubts and uncertainty in the process of layer definition (Lucas 2001, 154; Roskams 2001, 112). Overall, uncertainty claims describe interpretative doubts. For example, whether the extension of a layer is unclear. Fluidity claims include reports of reconsideration of previous interpretations. For example, changes in the interpretation of a type of feature. Possibility claims included descriptions when two or more interpretations are conceived to be plausible. For example, if it's ambiguous whether a layer was naturally or culturally deposited. Interactivity refers to doubts and reconsiderations derived from exchanges between fieldworkers and laboratory specialists onsite (Hodder 2005, 649). Evidently, this information is very similar to former types, but it was separated into a special type to examine in more detail the effects of this mechanism. Finally, methodological reflexivity includes comments about the methods and techniques implemented onsite. The last two types of reflexive claims were exclusively defined for reflexive diaries.

Low level claims: Contexts descriptions and finds reports

Current traditions of stratigraphic excavation in Britain define contexts and finds as the most basic entities to be defined onsite (Roskams 2001, 110; Carver 2009, 20). Contexts or stratigraphic units include deposition and removal events produced by cultural and natural agents (Harris, 1989, pp. 42–44). Depositional events are commonly known as layers, whereas removal events are defined as cuts. The elementary character of contexts derived from the idea that archaeological sites are formed by a sequence of events, hence the basic purpose of excavation is dismantling the site into its primary components (Museum of London 1994, 1.2). Gavin Lucas explains that contexts have a double identity as objects and events, however he observes that they are initially defined as objects (Lucas 2001, 159). Similarly, Martin Carver explains that interpreting a contexts means defining some sort of continuity of homogenous material or making a bunch of mud and stones into a layer (Carver 2009, 20). Hence, the first step of context definition requires identifying a stratigraphic unit which then must be described by giving account of its main physical attributes (Roskams 2001, 169).

One of the characteristics of British methods is defining cultural layers, this implies that archaeological remains such as walls can be sliced into different contexts like a makeup brick layer and a surface of plaster. Likewise, skeletal remains and coffins from burials can be defined as depositional events (Roskams, 2000, chapter 11). Overall, contexts aim to represent real depositional units, unlike arbitrary-level techniques that are considered artificial (Harris, 1989, p. 20). However, arbitrary-levels or spits (as British call them) can be introduced as complementary control mechanisms. Then a deep layer can be sub-divided into multiple arbitrary levels and each level will be identified as a context. Thereby, context might be used to represent natural and cultural events but also arbitrary levels of recovery (Roskams 2001, 230; Lucas 2001, 163).

Finds are the second basic component defined in British methods. Finds are commonly known as artefacts which include tools and many forms of discards and organic residues such as bones and seeds (Carver 2009, 20). Finds are conceived as one component of layer but unlike most soil particles that are discarded, finds are preserved for further study. Furthermore, finds are collected and bagged accordingly to different criteria such as fabric and recovery method. Yet, they are always recorded in relation to contexts, then the finds from a layer are separated from the finds of another (Figure 2.3). Sometimes additional recovery standard are required as separating finds from the top and the bottom of a fill or a three-dimensional recording for a special finds (Roskams 2001, 218–226). Excavation work also includes the recovery of different types of samples which also must be recorded in relation to a context of provenance, even if additional recording standards are implemented. Based in these notions five types of low-level claims were defined. Initially, this included context descriptions and descriptions of finds recovery. However, during the process of coding three additional types of claims were inductively defined. These include, Low-level contrasts or comparations among layers or finds, <u>expected absence claims</u> or claims when an expected attribute of a layer or a find was not observed. And finally, <u>indicative absence claims</u> that describe situations when an unobserved attribute or finds was considered evidence of something (Table 2.2).

Figure 2.3. The recovery of finds from a fill and the definition of a cut. Burdale, Yorkshire. (Copyright Julian Richards CC BY 4.0). https://doi.org/10.5284/1021540.

Stratigraphic claims: layer contacts, stratigraphy and correlations

One of the main goals of British stratigraphic excavation is the removal of contexts in the inverse order of deposition for defining the depositional sequence of contexts (Harris 1989, 138; Roskams 2001, 110). This is achieved by the study of the physical relationships among contexts, namely whether a context overlays, cuts or abuts another. Then, if one determines that a layer overlies another, then one can interpret that a unit is later than another. Likewise, if a deposit has been truncated by a cut, it will be inferred that the layer is earlier than the cut. Thereby, British methods separate two types of stratigraphic information, namely physical relationships among layers and the stratigraphy of a site which describes the depositional sequence of contexts. Hence, the sequence of a site is derived from stratigraphic relations (earlier/later), which in turn have been derived from the physical relationships (overlying/underlying) among layers (Roskams 2001, 154–155; Harris 1989, 152; Carver 2009, 20). Then, an important difference is that layer contacts are relatively observable (Figure 2.3), whereas stratigraphic relations and the sequence are always interpretations of unobservable phenomena from he past.

One of the main challenges of this process is identifying the most relevant layer contacts for the interpretation of stratigraphy, because a layer can have multiple contacts but only a few will be relevant for defining its stratigraphic relations. For example, a cut might be truncating many deposits and being filled by many layers, but only the deepest layer being cut and the deepest fill will be relevant for defining its most immediate earlier and later deposits (Museum of London 1994, 1.2; Harris 1989, 34). Finally, it should be said that stratigraphic relations are recorded in a very distinctive diagram, the matrix Harris (Figure 2.4). Whereas physical relationships might be recorded in sections or descriptively (Harris 1989, 36; Roskams 2001, 156).

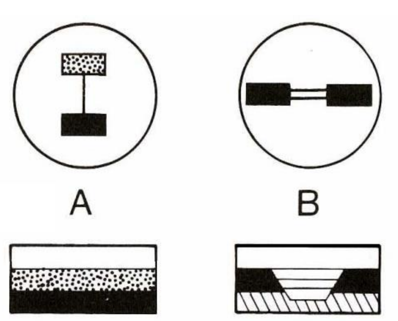


Figure 2.4. Representation of stratigraphic relations (A) and correlations (B) (above) and layer contacts (below). Source: Harris, 1989: 36.

Correlations are another type of stratigraphic data. These claims describe cases when two or more truncated contexts are interpreted as once forming the same unit. Correlations are derived from different evidence as the physical similarities among contexts or a similar position in the sequence, hence some archaeologist consider them to be slightly more complex stratigraphic relations (Roskams 2001, 156). However, correlations are interpretations of past processes too and therefore they are recorded in sequence diagrams with their own code of representation (Figure 2.4). In this way, stratigraphic information includes three types of claims: layer contacts, stratigraphy claims and correlations (Table 2.2).

Assemblage: group of finds

The assemblage is one of the most widespread notion in archaeology and for the same reason has a less specific meaning. Broadly speaking, an assemblage is a collection of finds defined due its provenience. Then, an assemblage can be a group of finds, from a site, a trench or a layer. For instance, Gordon Childe defined the assemblage as a representative collection from a site (Childe 1956, 16). But in modern British excavation, an assemblage commonly refers to the group of finds recovered from a context, therefore being equivalent to the notion of the finds (Roskams 2001, 212; Carver 2009, 224). But at the same time, there is an alternative notion of the assemblage as a group of finds with a patterned arrangement, burials are typical examples of this (Joyce and Pollard 2010, 293). This notion has some similarities with the British concept of the 'structured deposit' which presumes some sort of intentional arrangement in the depositional process of material culture. Generally, structured deposits are considered evidence of ritual practices (Garrow 2012, 94), however a similar idea has developed in other archaeological traditions. For example, Mesoamerican archaeologists usually consider that the distribution of goods in earth-offerings and burials are intentional arrangements that represent ancient cosmologies. In Britain, the concept of the structured deposition has gained popularity in last decades, but it has been noticed that it fails to consider instances of unintentional patterning (Joyce and Pollard 2010, 293; Garrow 2012, 91). For example, distributions of discard and tools over habitational floors (Flannery 1976, 34; Enloe and Audouze 2010, 18). Independently of its intentionality: burials, offerings and spreads of discard, usually require a careful spatial recording of the location and distribution of finds. This standard rests in the assumption that such patterning is evidence of something else. For these reasons, the assemblage was defined as a middle-level claim that escribes a patterned group of finds with detailed attention to its spatial attributes (Table 2.2).

Structural elements

The notion of the structure cuts across different concepts such as the layer, the feature and the group of features. For instance, there can be structural layers as walls, timber posts and floors. However, many of those structural layers are components of structural features, such as postholes and kilns. Finally, some groups of features define some types of structures like a fence or a building. Structural concepts are partially incorporated in SCR because there are special context sheets for the recording of masonry and timber layers (Museum of London 1994, 3.3-3.4). Moreover, these categories produced some conflicts with Carver's hierarchical system which only defines structures as groups of features (Carver 2009, 21–22). In other words, Carver's model doesn't consider that existence of structural layers and structural features. These distinctions are crucial because defining a structure is a characterization of a type of vestige, whereas defining contexts, features and groups of features are sets derived from stratigraphic units. In consequence, structural claims were organized as follow: structural elements for descriptions of structural layers and structural features (Table 2.2). Whereas the descriptions of structural groups of features and buildings were classified into higher-order categories.

Figure 2.5. Defining a brick floor. (Copyright Oxford Archaeology CC BY 4.0). https://doi.org/10.5284/1047571

Features: group of contexts

Features are commonly characterized as groups of contexts functionally related, therefore, these are considered a secondary claim that describes a high-order entity in relation to layers (Carver 2009, 20). A typical example of a feature is a human burial, this includes at least three contexts: a cut, a skeleton and a fill. Martin Carver classifies features into positive and negative, positive features grow up such as a kiln or a midden, whereas negative features include a cut that later is filled, for example a posthole (Figure 2.6). However, during the analysis it became more evident the features can be divided into structural and non-structural. Structural features include masonry and timberwork elements such as postholes and wall foundations, whereas non-structural features include everything else, such as middens and rubbish pits. In some extent, this coincides with the American definition of the feature that encompasses architectural elements such as hearths and storage pits (Flannery 1976, 5); yet, the American definition doesn't have the British stratigraphic background. For coding purposes, feature claims included descriptions of non-strcutural groups of contexts whereas structural features are included in the category of structural elements. This decision is based in the conceptual overlapping between structures and features (Table 2.2).

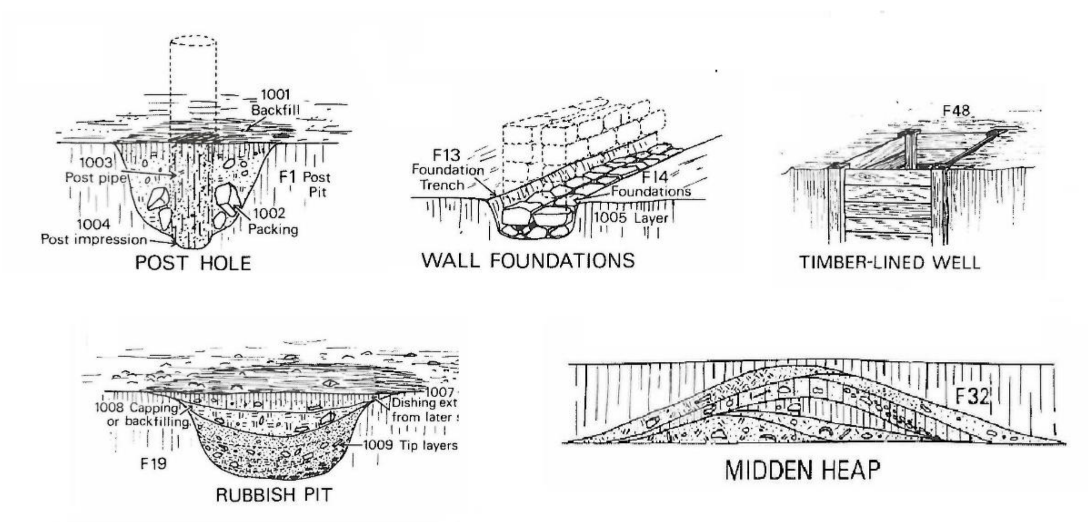


Figure 2.6. Types of features. Source: Carver, 2009: 21.

Group of features, Spaces, Buildings and Middle-level contrasts

As explained before, Martin Carver places the group of features as the following level of interpretation which defines structures such as fence lines and buildings (Carver 2009, 22). For example, a group of postholes defining a Mesolithic house (Figure 1.6). However, Flannery explains that a group of features also can define a space or an activity area (Flannery 1976, 34). This idea is integrated in Catalhöyük as they use the concept of the space to identify rooms defined by the internal walls of buildings (Figure 2.7). Likewise, they define external areas or alleys defined by the spaces between buildings (CRP 2014, 19). In this way, a group of features can define buildings and spaces at the same time, because architecture usually organizes cultural space (Zubrow,Audouze and Enloe 2010, 105).

Figure 2.7. Defining spaces of Building 77, Catalhöyük. (Copyright. Catalhöyük Research Project. CC BY-NC-SA 2.0).

Nevertheless, further conceptual revision also indicated that some groups of features can define areas and spaces but without defining a building. For example, a concentration of burials defines a funerary space and a concertation of storing pits can define a storage area. Flannery defines these as external activity areas (Figure 2.8). However, the definition of activity area is possible by the evidence of tool assemblages and discard scatterings (Flannery 1976, 34). A similar approach is adopted by French prehistorians that often interpret activity areas by studying the distribution of assemblages and the relation between features and assemblages, for instance, a hearth surrounded by bone discard. This strategy is highly common in sites that lack architecture or with minimal structural presence (Zubrow,Audouze and Enloe 2010, 106). Nevertheless, Flannery also incorporates the study of assemblages in urban sites to make finer spatial distinctions inside rooms and buildings the he defines as internal activity areas. For example, preparation and consumption areas.

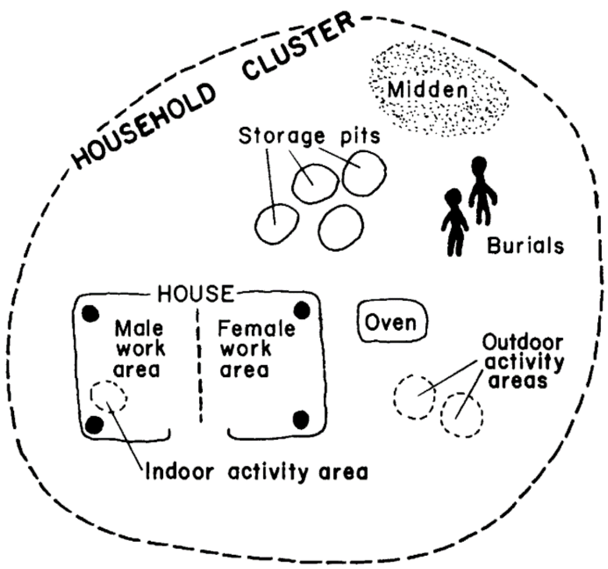


Figure 2.8. Types of activity areas. Source. Flannery, 1976: 45.

Overall, the group of features, the space and the building are closely related concepts, although they are not the same. For example, some groups of features define spaces and buildings, but some group of features define spaces. However, not every spatial interpretation will be derived from architectural elements. But in general, group of features, spatial and building interpretations are grounded in lower-level evidence. This revision of spatial concepts is not exhaustive as there can be further levels of interpretation, for instance groups of buildings, but since such types of claims were not observed in the samples, they are not considered then. In sum, three additional types of claims were defined. Group of features for any kind of set relation that includes two or more features. Spatial claims for different types of spatial interpretation like rooms and storage areas and finally building claims a special type of structure defined by a group of features and spaces, for instance a house. The analysis allowed to identify an additional type of claims when testimonies describe similarities and differences among two middle-level entities, these were defined as middle-level contrasts (Table 2.2).

Formation processes

After having defined a layer and additional interpretative step is the investigation of formation processes (Roskams 2001, 169). In broad terms, formation processes aim to explain the depositional process of contexts or how layers came to be in the first place and the subsequent post-depositional transformations (Schiffer 1972, 156). The general notion is grounded in the fact that archaeological sites are product of cultural and natural agents. Then, the purpose of formation process is to define whether a fill was deposited by a water-current (natural layer) or if it was dumped in there (cultural deposit). In some cases, the distinction between natural and cultural events might be relatively simple as with structural layers. Likewise, human cuts and animal burrows are also relatively simple to differentiate. Yet, formation processes are always inferred from physical attributes of layers and finds. In other words, archaeologists do not observe formation processes (Schiffer 1987, 265).

Michael Schiffer has proposed a complex terminology for the description of cultural formation processes (Schiffer 1972, 161–162). However, Gavin Lucas has observed that much of Schiffer's work focuses in the depositional processes of artefacts and assemblages (Lucas 2012b, 75). The analysis of records showed the regular use of an alternative terminology to describe the formation process of strcutural remains which commonly differentiate between construction, refurbishment, use and destruction events. For example, cuts, walls and floor-patches can represent manufacture and refurbishment. A worn-out floor might represent a use event, and finally a collapsed wall represented a destruction event. On the contrary, the interpretation of formation processes in sedimentary layers is grounded in a more general distinction between primary and secondary layers. Primary layers usually describe layers found in its original place of deposition, for instance a charcoal fill inside a hearth or an undisturbed skeleton in a burial, whereas a secondary layer implies a relocated deposit, for instance a dump or a relocated skeleton. Finally, post-depositional transformations usually were described as truncations or as natural disturbances.

Lucas has claimed that single context recording focuses in layer description and stratigraphy, whereas the interpretation of formation processes is not part of the system (Lucas 2012b, 88). This is a more specific claim of the general idea about the limited interpretive scope of fieldworkers. Moreover, this idea is partially subscribed by Harris because he considers that the primary task of fieldworkers is interpreting the sequence not formation processes (Harris 1989, 53). Nevertheless, formation process theory is compatible with stratigraphic theory as both acknowledge the existence of natural and cultural events. In consequence, the challenge of formation process interpretation is investigating the specific types of depositional process for each layer, with an emphasis in cultural events.

Phases

Phases are chronological division of the stratigraphic sequence. Then each phase represents a group of contexts chronologically related or relatively contemporary (Harris 1989, 158). For example, a constructive phase of a building or an occupation phase of a site. Specifically, single context recording prescribes that phasing must be done offsite as part of the post-excavation program (Roskams 2001, 257; Spence 1993, 27). However, some preliminary phasing might be possible in not very complex sites, moreover some alternative field methods are keener to start interpreting phases onsite (Harris 1989, 108–115). For this reason, an initial consideration of phase claims was included. Stratigraphy, formation process, phase claims are similar as all of them describe different types of depositional processes. However, formation processes are specific for each layer, whereas stratigraphy links individual events in a sequence. Meanwhile, phases describe a chronological sequence arranged in groups of layers, features and structures.

High level claims: Function, Agency and Culture/Period

As commonly accepted, the main goal of archaeology is the interpretation of the past. However, this starts very early in excavation work with the interpretation of the stratigraphic sequence and formation processes. Then, a more precise claim is affirming that the main aim of archaeology is the interpretation of the sociocultural past. Hence, these types of claims constitute the higher-level of interpretation. After a preliminary coding of records, three types of high-level claims were found. Functional, agency and culture/period claims. Originally, an additional type was included for symbolic claims, but no instances were found, then being discarded. Functional claims include interpretation about the use of remains for certain tasks. For example, the use of a pit for dumping rubbish or the use of a space for storage. This characterization is inspired in processual theory that describes cultural groups as systems formed by several subsystems with specific functions for the maintenance of the system (Johnson 2010, 73). Hence, there is a similar logic in the functional interpretation of sites which are conceived as mechanisms where every feature, structure and space has a specific purpose or a set of tasks to achieve.

In relation to agency, Hodder explains there are different concepts. However, a compatible notion with information observed in records is the characterization of agency as individual action and intentionality (Hodder and Hutson 2003, 101). This definition coincides with a recent call for the sociocultural interpretation of stratigraphy which suggest that layer interpretation also should include the cultural actions behind depositional processes, for example instead of just interpreting a layer as a dump, one should ponder whether this could be an instance of a cleaning action (McAnany and Hodder 2009, 1). Finally, Culture/Period claims represented instances when fieldworkers interpret material culture in relation to a culture or a period. For instance, if they define a feature as a roman kiln or medieval ditch. Obviously, these types of claims are related to the notions of Culture History (Johnson 2010, 15).

In general, high-level claims depend on the evidence of low and middle level claims. For example, the functional interpretation of a feature commonly derives from layer and feature attributes, and the spatial relation with other features. Likewise, the formal attributes of artefacts and structures can be indications of their cultural date. Similarly, as Schiffer observed long ago, the investigation of formation processes would allow to produce more reliable inferences about behavioural claims (Schiffer 1972, 163). In consequence, functional, agency and culture/period claims are sociocultural claims of remains (layers, finds, structures and so forth) and depositional processes (stratigraphy, formation processes and phases). Evidently, there are other additional forms of high-level interpretation like the site-interpretation and historical synthesis (Carver 2009, 298).

Table 2.2 Classificatory framework for excavation records.

Class/Level	Type of claim	Brief description
Non-Interpretative	Circumstances	Descriptions of observational conditions related to excavation work. For example, a layer being dug under artificial light
	Decisions	Descriptions of operative decisions related to excavation work. For instance, a layer being dug in arbitrary levels.
	Ethnographic	Descriptions of activities related to the project but not interpretative work. For example, having visitors onsite.
Reflexivity	Uncertainty	Interpretive doubts. For example, uncertainty regarding the extension of layer.
	Fluidity	Interpretative reconsiderations. For example, changes in the functional interpretation of a feature.
	Possibility	When two or more interpretation are plausible. For example, when it’s unclear whether a layer was naturally or culturally deposited.
	Interactivity	Interpretative ambiguities and reconsiderations produced by the interaction among fieldworkers and specialists.
	Methodological	Discussions about of methods and techniques in primary records.
Low-level claims	Contexts	Layer and cut descriptions.
	Finds and samples	Claims about the presence/absence of finds and the recovery of samples.
	Low-level contrasts	Claims about differences and similarities among contexts and finds.

		Expected absences	Describing the absence of expected context attributes or finds.
		Indicative Absences	When an absent layer attribute or find was indicative of something else.
Middle-level claims	Stratigraphic claims	Layer contacts	Physical relationships among layers: overlying, cutting, abutting.
		Stratigraphy	Stratigraphic relation and sequence claims, for example if a context is earlier/later than another.
		Correlations	Descriptions of two or more truncated contexts that once were the same unit.
	Middle level entities	Assemblage	Descriptions of a group of finds that includes spatial attributes. For example, a burial.
		Structural elements	Descriptions of structural layers (walls, floors) and structural features (hearth, kiln, platform).
		Features	Descriptions of a group of contexts. For instance, postholes, ditches and middens.
		Group of features	Any description of two or more features associated into a group. For example, a line of postholes defining a fence line.
		Spaces	Descriptions of internal or external areas. For example, a room or a funerary space
		Building	A type of structure defined by a group of features and spaces. For example, a house.
		Middle-level contrasts	Descriptions of similarities and differences among middle-level entities.
	Middle level	Formation processes	Descriptions about the natural and cultural depositional processes of contexts. Including post-depositional transformations.

		Phase	Claims related to the definition of constructive and/or occupational phases of sites.
High level claims		Functional	Claims about the use of remains for certain tasks. For example, the use of a pit for dumping rubbish or the use of a space for storage.
		Agency	Descriptions of past actions in terms of individual/group intentional actions. For instance, cleaning a structure or destroying a building.
		Culture/Date	Interpretations of remains in relation to a culture/period. For instance, the definition of a roman kiln or a Neolithic burial

Chapter 3 Reflexive archaeology

A *notebook of experiments is not to be confused with moment-by-moment diary of everything that happened in the laboratory. Marc Bloch, 1953, 120.*

The Catalhöyük Research Project is the one projects where reflexive methods have been put into practice. This chapter describes the results of the internal analysis for this case study, particularly to explain the positive and negative aspects of diary recording for excavation recording. Initially, this chapter provides a general description of the reflexive method with an emphasis in the role of diaries and site-tours (3.1). This is followed by a description of the epistemic circumstances of the project and the character of the site (3.2). Reflexive diaries are stored in the digital database of the project (www.catalhoyuk.com/research). However, due the size of the project and the organization of explorations, the analysis required the selection of a specific building (B49) for the analysis. Hence, the chapter includes a brief explanation of sampling criteria (3.3). The following step consists in providing a general characterization of the format of diary entries (3.4). The main section presents the qualitative and quantitative results of the internal analysis (3.5). These results lead to a revision of the notion of reflexivity which will allow to differentiate between two forms of reflexive testimony: observational and evidential. This distinction will contribute to assess the relevancy of reflexive data.

At this point, there will be enough information to provide a general description of interpretative performance at B49. Specifically, to evaluate the impact of additional recording tools for improving the credibility of field interpretations and the effect of having lab specialist onsite (3.6). The analysis is completed examining the individual performance of diggers (3.7) and with a longitudinal assessment of the exploration at B49 (3.8). These elements will be useful to explain how individual skill and the accumulation of background knowledge affected the interpretative and reflexive process at B49. The chapter concludes resuming the main positive and controversial aspects of diary recording at Catalhöyük. Hence, this evaluation considers methodological elements and specific epistemic circumstances of the Catalhöyük Research Project (Section 3.9).

3.1 The reflexive method

The reflexive method is part of a wider theoretical program called reflexive archaeology which has been described as a methodological and ethical reaction against currents field practice (Hodder 2005, 644). Broadly speaking, ethical aspects assess the relation between archaeologists and other stakeholders like indigenous groups. Whilst, the methodological reaction is mainly against the procedures of commercial archaeology. Overall, the reflexive program is based in four principles. Reflexivity, which is defined as the capacity of examining archaeological assumptions. Secondly is contextuality, which claims that archaeological evidence gains meaning by considering their contextual relations. Thirdly, is interactivity, which emphasises the necessity to develop collaborative mechanisms to question and criticize interpretations. Finally, there is multivocality which seeks to create space for multiple discourses about the past (Hodder 2005, 648). As it will be explained, the first three elements are closely connected to field methodology whereas multivocality is closer to ethics, therefore it shall not be considered here.

Hodder believes that reflexivity, interactivity and contextuality are angular stones of interpretative work, but these had been minimized due standardized protocols in commercial projects. For example, he suggested that reflexivity is restrained by standardization that portrays data as fully objective and ignores doubts and ambiguities (Hodder 1999, 31). Likewise, contextuality and interactivity are constrained by disciplinary strong separations between fieldwork and post-excavation, then creating an false divisions between description and interpretation that rarely achieved an integrated view of archaeology (Hodder 2005, 645). Hence, the reflexive tools and mechanism aim to encourage more reflexivity, contextuality and interactivity. For example, video and diary recording aim encourage fieldworkers to be more critical about their own work. Similarly, priority tours aim to create a closer relation between fieldworkers and laboratory specialist that enable exchanging data and revising interpretations. There are other tools and mechanism in reflexive diaries, however diaries and priority tours have a stronger incidence in excavation work. Moreover, the reflexive method included many digital technologies with the goal of improving interpretative practice in its later years (Taylor et al. 2018; Berggren et al. 2015). However, these aspects do not form part of the analysis.

Diary recording

Reflexive diaries are one of the most characteristics elements of reflexive methods. However, it should be very clear that diaries represent a complementary format in addition to standardized sheets. This is an important element to avoid confusions between formats (context sheets, diaries) and methods (SCR, reflexive method), which depend in the way that the use and integrate formats in a recording strategy (see section 2.2). Specifically, reflexive diaries constitute an additional sequence of textual recording on top of sequences of context and feature sheets or as Hodder mentions:

> *"The excavator still works scientifically and systematically, filling inn forms and taking samples and grid coordinates. But in addition, the reflexive excavator is also doing other things -talking to specialists and community members, writing diaries, making videos, recording his or her phenomenological experience and so on"* (Hodder, 2005, 651).

As mentioned before, reflexive diaries aim to provide a complementary information that apparently standardized methods fail to include, for example decisions, ambiguities and reconsiderations (Hodder 1999, 31). Then, reflexive archaeologists assume that having an additional layer of recording will contribute to improve the credibility of records by capturing crucial contextual data omitted in more traditional methods (Hodder 2000a, 9).

Reflexive diaries are inspired in the old excavation notebook, but instead they are captured in an electronic database onsite at the end of a day's work. More importantly, there have been important changes in the procedures of dairy recording along the development of the project. For instance, reflexive diaries aimed to be less codified and less restrictive in the early years of the project, to facilitate the description of a more personal view of the site and develop higher orders of interpretation (Farid 2000, 25). However, a different attitude emerged in subsequent years because it was noticed that unexperienced fieldworkers often recorded irrelevant data due the lack of guidelines (Farid 2015, 73). Likewise, it was noticed that fieldworkers were not very systematic with diary recording and this required a stronger policy to regulate recording activities (Berggren et al. 2015, 436–437; Berggren and Nilsson 2014, 62). Thereby, the last instructive of site procedures defines clearer objective for diary recording. For example, describing the process of interpretation at different scales (layer, feature, space) and being reflexive about the effects of methods in interpretative practice (CRP 2014, 33). Likewise it was expected that fieldworkers could make one or two entries a week following a narrative structure (Hodder 1999, 95; Mickel 2015, 305). Additionally, there have been other important operative changes in diary recording over time. Specifically, diaries were exclusively recorded by trench supervisors in the early years of the CRP, however this task was extended to most excavators later (Hodder 1997, 696). This had the intention of giving more interpretative chances to diggers. However, this was not evenly welcomed as some people considered it more like an additional burden which compromised having a basic systematic record (Farid 2000, 25; Berggren and Nilsson 2014, 62).

Priority tours

One of the most distinctive characteristics of the reflexive method is having a group of laboratory specialist onsite running preliminary analysis of samples and artefacts and feedbacking results to fieldworkers (Figure 1.8). The interaction among fieldworkers and laboratory specialist is regulated by ‘priority tours’ in which a group of lab specialists visit excavations for selecting samples and discussing results with fieldworkers (Farid 2000, 19). Thereby, priority tours develop a recursive loop that puts together a wide range of data that could either reinforce or contradict a preliminary field interpretation (Hodder 2005, 648). Sometimes, the process of sampling is called ‘prioritizing’ or ‘fast-tracking’ (Hodder 2005, 653; Farid 2015, 68), but more importantly this is defined as a negotiation among archaeologists, because instead of following a predetermined policy, they select samples considering a wider range of criteria. Then, an specific goal of reflexive diaries is giving account of sampling decisions and explain the implications of laboratory results for field interpretations (CRP 2014, 33).

In recent years, there has been some controversies about the originality of reflexive methods because many of them have previous versions (Carver 2011, 25; Thorpe 2012, 33; Hanson and Rahtz 1988; McAdam 1992). More recently, the reflexive method has been describes as the first attempt to systematize such mechanisms (Berggren 2015, 6250). Finally, despite its association with postmodern philosophies, reflexive principles are grounded in two very old methodological ideas. Specifically, reflexivity predicates that an unexamined field-practice is not worth doing and therefore, they promote a critique of one’s practice and an evaluation of opinion by dialogue as suggested by maieutic. This is not the place to follow such treads. For now, it suffices to say that reflexive method is a set of mechanism that seeks to improve interpretative practice.

3.2 The project and the site

The first case study revises the excavations of the Catalhöyük Research Project (CRP) in the famous Neolithic site of Catalhöyük, in Turkey. The site is constituted by twenty meters mound that includes a deep sequence of mudbrick buildings, constructed and dismantled one above the other representing 1400 years of occupation. The site is famously known for its well-preserved buildings that include wall painting, relief sculptures and architectonic installation with animal remains. Some areas of the site include roman and byzantine remains but the main occupation dates from the Neolithic (Farid 2015, 60). Before the CRP, there was another important investigation in the site by James Mellaart in the 1960's, those excavations provided a general description of the architecture of the site and an initial interpretation of the function of buildings. Specifically, Mellaart suggested buildings included houses and shrines (Mellaart 1967, 3).

The CRP began in 1993 with an evaluation phase, followed by three phases of excavation between 1995-2017. So far, only the first two have been published covering results between 1995-1999 and 2000-2008 (Hodder 2006, 2014). This long-term research project had explorations in the East and West mound, although the main exploration were located in the former (Figure 3.1). The North and South areas were the main excavations of the project, the South was a deep exploration for investigating the constructive sequence of the site, whilst the North operation was a big horizontal exploration to investigate the organization of the settlement. Initially, the North exploration began with two building (B1 and B5) to investigate their function, suggesting a domestic use. In 2003, the North area was expanded to explore a larger number of buildings (Hodder and Farid 2014, 3). It's in this phase of excavation that B49 was dug (Figure 3.2). Moreover, this phase was correlated with operative changes in diary recording, because every fieldworker onsite should make this record, instead of site-supervisor only as in previous years (Berggren and Nilsson 2014, 62).

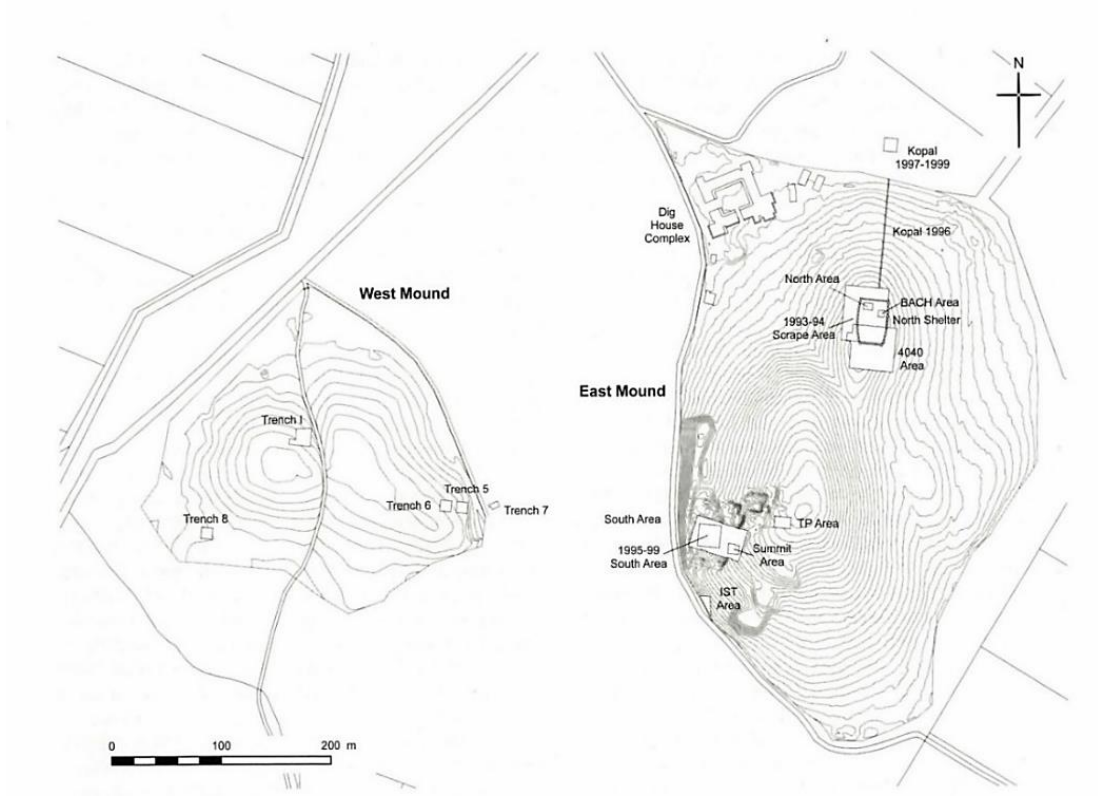


Figure 3.1 Excavations at East and West mound of the Catalhöyük Research Project (Copyright Catalhöyük Research Project).

Ian Hodder, the famous scholar behind the theorical ideas of reflexive method, also was the director of the CRP. Although, fieldwork supervision was mostly in charge of Shahina Farid, an archaeologist emerged from the ranks of commercial archaeology. She fulfilled this role between 1995 and 2011. The project included the participation of many international teams, but the British team had the leading role as it controlled the largest explorations (North and South) and had the longest participation. The British team was formed by a group of professional diggers from the commercial sector especially hired for this purpose (Farid 2015, 76). However, the British areas also included the participation of scholars and trainees from American universities like Stanford. Among the international teams, there was an American team from UC Berkeley University in charge of BACH area, a Polish team for the TP area and a Turkish team for the IST area to name a few. These teams usually had shorter participation and controlled smaller excavation areas, and often they followed a mix of British and their own methods. Besides, they usually included a mix of scholars and trainees (Farid 2015, 67; Trigham and Stevanovic 2000).. The project was funded by different types of sponsors from academic and government institutions, including private companies. Among other things, this allowed the construction of the 'dig house' a building facility with a visitors centre, laboratories, storage and accommodation areas located in the margins of the site (Hodder and Farid 2014, 2).

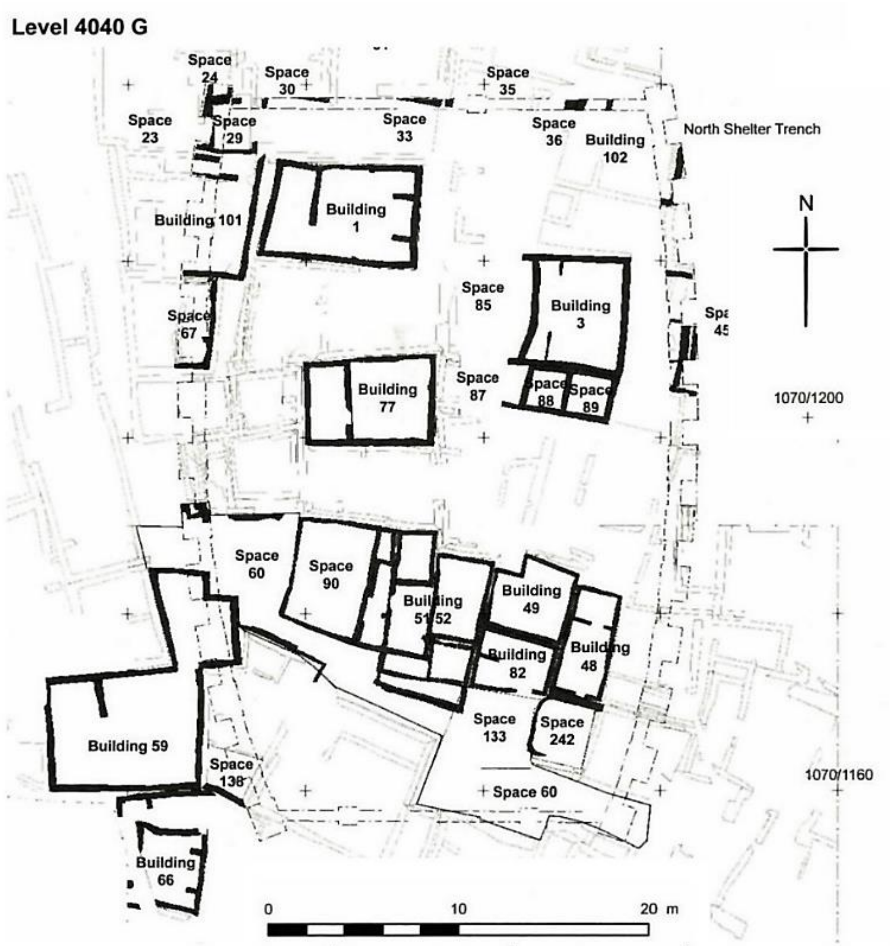


Figure 3.2 The North area (Copyright Catalhöyük Research Project).

Building 49 (B49)

Given, the scale of explorations and the number of participants, the CRP has been described as a large-scale reflexive experiment. For this reason, the exploration of B49 was the selected to investigate the implementation of reflexive procedures. B49 was a small squared-like building of approximately four meters by its sides. It included a main room (Space 100) and a smaller side room (Space 334). The south area of the main room had a concentration of hearths and ovens with 'dirtier' surfaces, whilst the northern half had a concentration of cleaner platforms containing human burials. Likewise, the northern half had internal wall painting (Figure 3.3).

After the exploration, B49 was defined as a domestic building with evidence of numerous architectural modifications but always respecting the original plan (Eddisford 2014, 313). Despite size differences, B49 was a very typical Neolithic building that follows the general pattern observed across the site (Figure 2.7).

Figure 3.3. The final constructive phase of B49 (Copyright Catalhöyük Research Project CC BY-NC 4.0).

B49 was identified by the end of 2003, almost immediately after clearing topsoil (Krotscheck 2003, 23) (Figure 3.4). The exploration continued in interleaved seasons in 2004, 2006 and 2008. Generally, explorations lasted four or five weeks each year. In 2004, B49 was explored by a small group of people from the Stanford field school, this included two graduate students supervising and four trainees. In 2006, two professional diggers continued the exploration. Finally, the excavation concluded in 2008 and one of the former professional diggers continued as supervisor assisted by five new participants, including another professional digger, a graduate student and an undergraduate student (Eddisford 2014, 356). Overall, B49 was excavated by twelve diggers organized in small units. Thereby, this account contrast with the common description of the project as large-scale reflexivity (Jones 2013, 180; Farid 2015, 59).

Figure 3.4. Detection of B49 (Copyright Catalhöyük Research Project CC BY-NC 4.0).

3.3 The sample

Catalhöyük diaries are stored in a huge online database (www.catalhoyuk.com/research) that includes hundreds of entries recorded between 1993-2017 and given the scale of the project, the most immediate task was defining a manageable sample. Specifically, the North area was selected for this purpose due its strategy and being under the management of British archaeologists. Besides, priority was given to the second phase of the project because in this period most fieldworkers participated in diary recording. Finally, since explorations were organized by building, one of these was selected to examine diaries. This selection was no less complicated because despite most building explorations had produced some entries, sequences included a few entries because recording had been unsystematic. Besides, some sequences only reflected the opinion of supervisors. These were negative aspects didn't contribute for the analytical strategy, which required a more systematic account and participation of various team members. No further detail of these drawbacks will be given as they have been described elsewhere (Berggren and Nilsson 2014, 62).

B49 was the best case for analysis because it was one of the fewest buildings systematically documented. Nonetheless, data collection had its own problems because access to diaries is provided by a web-searcher that offers predefined categories (year, author, day, building, space, feature and unit). After searching by building, the browser returned 17 entries from 3006 and

2008. However, when searching by space (100 and 334), the database released 27 additional entries from 2004. Finally, three additional entries were found when searching by author. Overall, sampling was complex due the scale of the project, but collection was complicated by archiving problems. The final sample included 47 entries, these were ordered in sequence by date and tagged with a code that contains the initials of the author, a sequence number and year of production. Thus, "*UK/01/04*" is the first entry of the B49 sequence, recorded by UK in 2004. The length of entries varies from a short paragraph to a few pages, but the average extension is one page. The total amount of analysed text involves 45 pages, which cover around 15,500 words. As explained before, no context sheets from the Catalhöyük archive were examined because there is a study already that describes similarities and differences in the content of sheets and diaries within the reflexive method (Mickel, 2015). But more importantly, the primary goal of my investigation consists in evaluating the relevancy of reflexive information for stablishing the credibility of data.

3.4 The narrative format of reflexive diaries

One the clearest characteristics of B49 entries is a predominant narrative format which usually relate different episodes of the act of discovery (Edgeworth's sense). These chronicles usually describe successive steps in the exploration of remains for instance, a layer, a structure or a building. In this way, B49's entries coincide with the general narrative format of reflexive diaries (Mickel 2015, 303). However, diaries are selective because there not a systematic record of every act of discovery. Another important characteristic is the textual format of diaries as none of these testimonies is supported with sketches. Then, a typical testimony follows this pattern:

> *Today we took out yet more room fill, and contrary to my claim from yesterday, we continued to fast-track. A thorough cleaning of the entire space for the priority tour clarified a lot of the features emerging underneath the fill that we had been taking out for ca. 5 days. We re-found and planned at least three basins or bins in the west part of space 100 (two rectangular, one round), and the beginnings of a plaster line crossing E-W in the northern part of the space, which might possibly be the top of a platform or a similar feature. (UK/05/04).*[2]

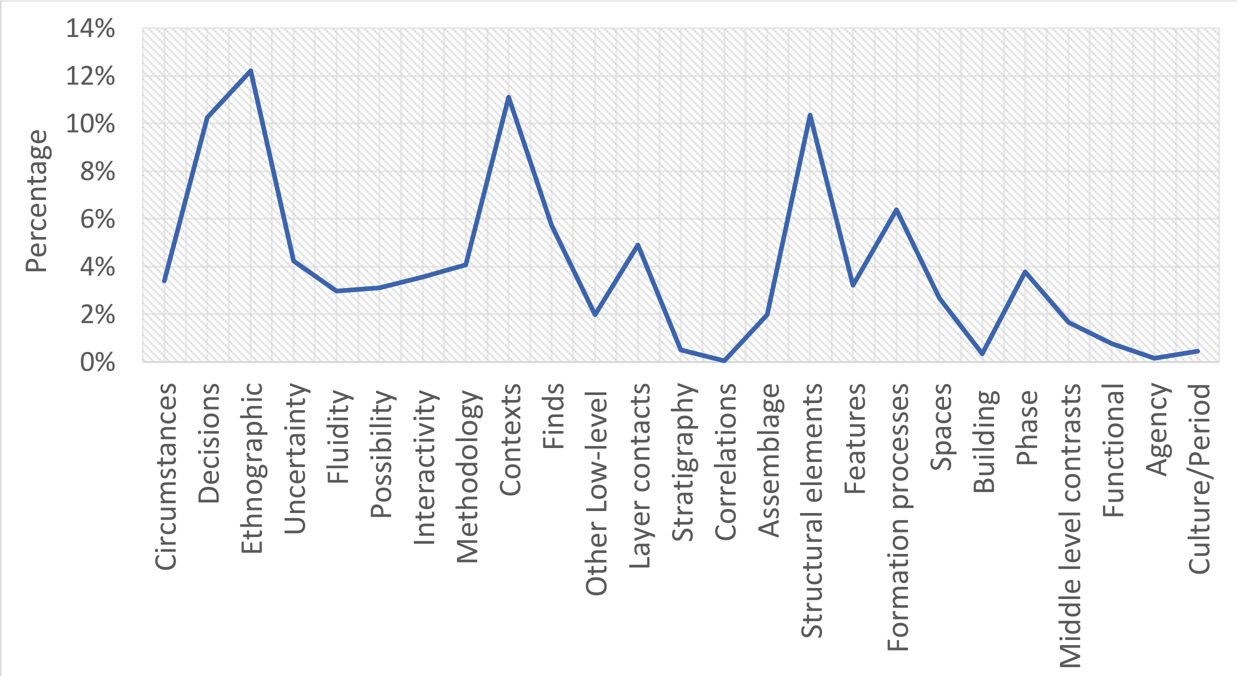
[2] An identification code is used here and onwards for referencing testimonies of B49's entries. The code indicates the initials of the author, the sequence number and the season when the entry was recorded.

Another general characteristic of entries is that diggers usually describe their personal observations, but at the same time they also report discoveries made by others. This indicates that diary recording also aims to give account of teamwork and is not limited to report individual practice. Sometimes two or more entries contained duplicated information recorded by different people. For instance, the fragment below replicates the information contained in the testimony above. Duplicated content included different types of claims and was observed in entries for 2004 and 2008. The quantity of duplicated information was not calculated, then if a claim was repeated, it was coded and counted twice.

> *Today was a good day for plaster. I worked near the north wall, taking out the rest of the room fill (unit 7913). This unit is still on fast track, even after several days of working with the same unit number. We successfully identified a platform along the north wall. I spent most of the day following the plaster to define the feature. It slopes quite steeply to the north. (SL/06/04).*

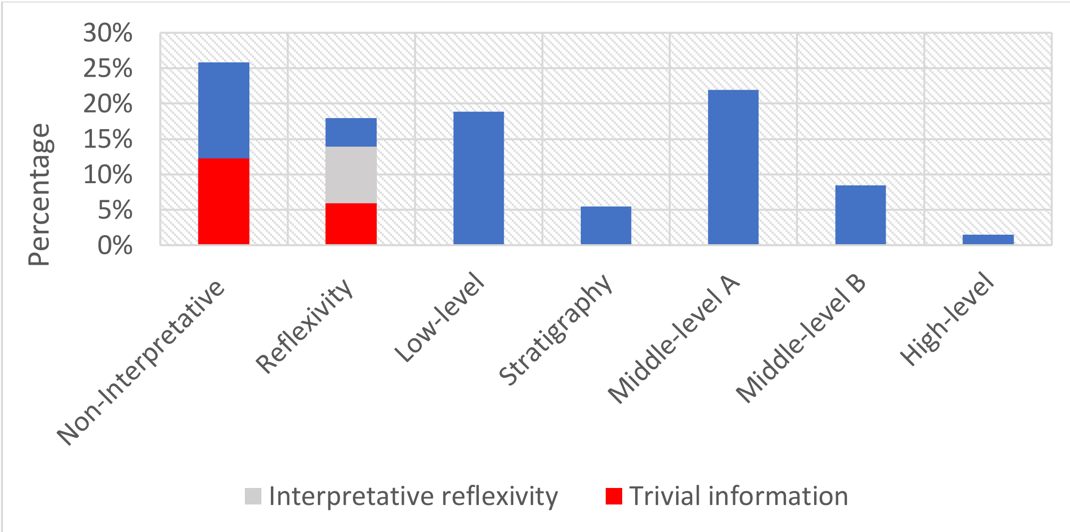
3.5 Results

The main purpose of internal analysis was to examine the distribution and the relevance of content in diaries. Graphic 3.1 shows the distribution of information by type. The most frequent are two types of non-interpretative claims (Decisions and Ethnographic claims), and two types of interpretative claims (Context and Strcutural Elements), which particularly relate the discovery of such vestiges. These data are consistent with the character of B49 which was largely constituted by structural layers and features. Less frequent are claims about layer contacts, features, formation processes and reflexive claims (Uncertainty, Fluidity, Possibility, Interactivity and Methodology). Yet, the lowest values are related to the highest-level of interpretation (Functional, Agency and Culture/Period). Most of these types are common with further cases studies, except for Ethnographic, Interactivity and Methodology that are exclusive of B49.



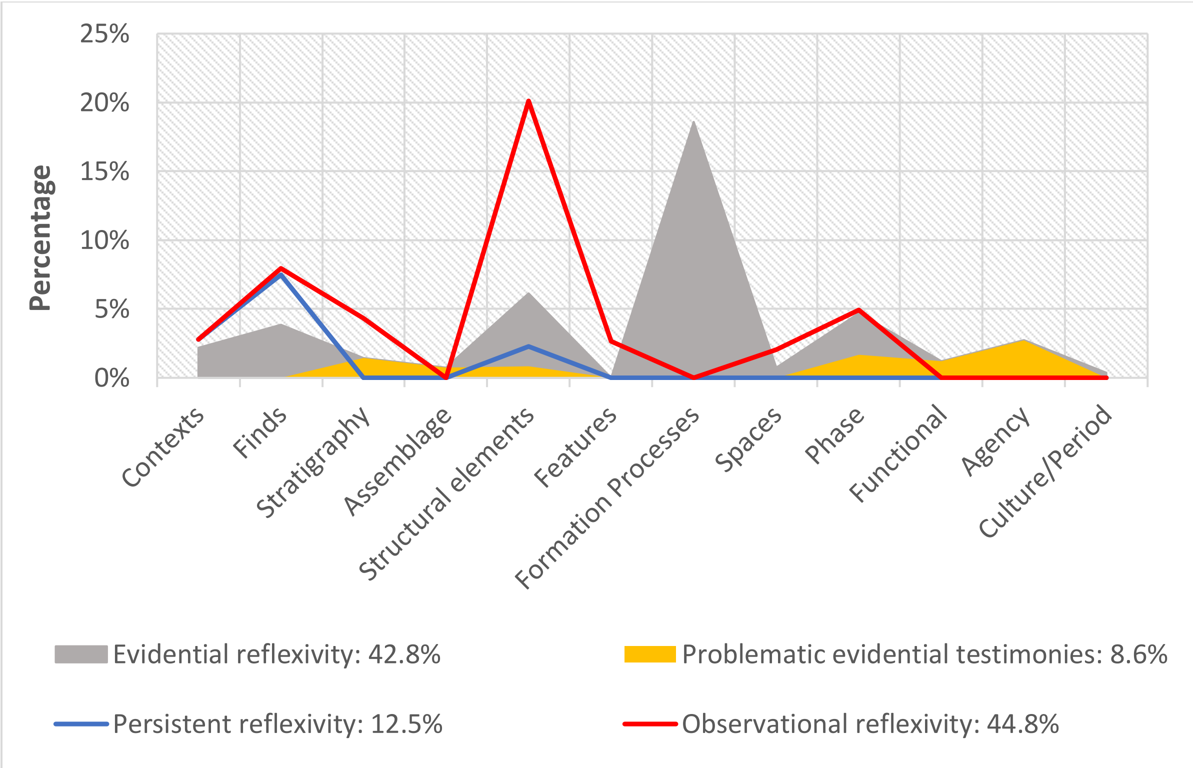
Graphic 3.1. General distribution by type. B49. The graphic illustrates the variability and percentage of information types in excavation diaries.

When organized by class, a quarter of the total information represents non-interpretative claims and almost a fifth covers reflexive claims (Graphic 3.2). Thereby, almost one half represents non-interpretative and reflexive claims, meanwhile the other half integrates interpretative data. One of the main results of the analysis is the identification of irrelevant reflexive and non-interpretative content in reflexive diaries (Graphic 3.2). However, the proportion of trivial reflexivity has been calculated in relation to the percentage of interpretative reflexivity. This includes all reflexivity types except Methodology.



Graphic 3.2. General distribution and relevancy by class. B49. This graphic illustrates the distribution of information by class/subclass and the percentage of trivial information for non-interpretative and reflexive classes. The proportion of trivial reflexivity has been calculated in relation to the percentage of interpretative reflexivity. This includes all reflexivity types except Methodology (blue area of the column). Middle level A summarizes assemblage, structural elements, features and formation processes claims. Middle level B summarizes building, space and phase claims. High-level summarizes functional, agency and culture/period claims.

One of the main results of the analysis was the classification of reflexive claims into three categories (Observational, Evidential and Persistent reflexivity). Graphic 3.3 illustrates the distribution of reflexive information according to these categories and in relation to diverse areas of interpretation. In general, there is an even proportion between observational and evidential reflexivity, although most of observational reflexivity is related to the investigation of structures and most of the evidential reflexivity is related to formation processes. These distinctions were very useful to differentiate between relevant and trivial data (see section 3.6). The graphic also includes the proportion of recording problems for evidential claims. A detailed description of these results will be the purpose of the following sections.



Graphic 3.3. Reflexivity B49. This graphic illustrates the dstribution of reflexive claims in relation to three qualitative categories (evidential, observational and persistent reflexivity) and interpetive types. The graphic also includes the proportion of problematic testimonies for evidential testimonies.

3.5.1 Non-Interpretative information

Non-interpretative claims include three types of information. Data about observational circumstances affecting excavation work was the less frequent (Graphic 3.1). This was a recurrent subject in 2004 when the site didn't have a shelter, and this caused some observational problems with operative implications.

> *The shelter is due to go up within the next few days and we have been instructed not to expose any of the plaster until it has. Also, DGN and I stopped working to remove the fill in the southwest corner because the visibility was just too bad to detect the subtleties in the soil colour change. (SL/02/04).*

The proportion of decision claims was higher, and this type includes testimonies of diverse forms of operative decisions like in the example above. However, even when working in ideal conditions technical difficulties arose, and diggers might report particular solutions for specific challenges like in the following example:

> *Since burial f. 4000, I brought down the NW platform, 1651, out of phase in order to create a more workable area, as the burials are now over half a meter down from the surface of the platform. There are some semi-articulated remains at the bottom of the cut that need to be investigated but have been put on hold until f. 1651 becomes more accessible. (CM/42/08).*

Some examples describe instances of recording procedures, for example when a layer had been dug in two interventions and each portion had been assigned a different context. Then diaries clarify these procedures.

> *Platform F4004, in the SW corner was partially excavated in 2004, when an ashy midden like construction core (7957) was removed and found to contain a cluster of animal figurines (7958). This unit was fully excavated this season as (14460), and contained a high concentration of finds, including worked and unworked stone, animal bone, worked bone, obsidian, red ochre (?) and an antler with the binding of the handle visible as phytolithis. (DE/36/08).*

Less frequent and more important are some examples of recording decisions with interpretive implications. For example, when not being sure about the extension of a layer and then arbitrary layers help to control the excavation (see examples in the next section). Thereby, operative decisions can be organized in two classes, those with interpretative implications and those without interpretive consequences. As explained before, one of the characteristics of the reflexive method is its sampling strategy based in ‘negotiations’. However, when diaries describe this process, testimonies generally fail to give account of decision criteria. Therefore, failing to achieve the methodological goals (Hodder 2005, 652; Farid 2000, 70). Specifically, diaries never discuss the relevance of factors like the abundance and type of finds or the interpretive potential of samples. This doesn’t mean that sampling decisions were not based in complex criteria, only that diaries do not report them. Still, diaries are not completely silent, for instance, the fragment below suggests that sampling responds to the primary character of the layer, probably a rich organic assemblage.

> *We prioritized 7948 to see how it would relate to other fills found on top of fire installations. We also plan to prioritize 7949 and 7951, which correspond to the fill in our ovens. The fine white layers we had located yesterday under 7931, concentrated in the corner near the north platform, were examined by the botany team. Emma took a sample (7931 S4) to see what the material might be. (UK/22/04).*

Finally, ethnographic information was the most frequent type of non-interpretive claims. These claims were commonly located at the beginning and/or the end of entries and describe different activities related to the project but unrelated to interpretative work, for example having visitors onsite or when someone had a break from excavation work among others.

> *First days of work in building 49 in the 4040 area. Have spent most of the day familiarising myself with the area and the recording system, as it is my first year at Catal. Tomorrow the real excavation work will begin and hopefully the floors will behave and go down the way they are supposed to.* (LF/32/08).

Another recurrent topic are the psychological and physical effects on diggers, for instance excitement, tiredness and stress. These testimonies represent the most personal information in diaries. In general, ethnographic testimonies describe some aspects of the social life of the project. In this way, non-interpretative claims represent different forms of contextual information. However, circumstantial and operative information represent relevant data to contextualize interpretative work, whereas ethnographic claims are irrelevant for this purpose (Graphic 3.2). This is a crucial element to consider because B49 diaries invest a lot of space on irrelevant aspects.

3.5.2 Contexts

Despite one of the main goals of reflexive diaries is transcending the scale of context definition, one of the most frequent types of account are descriptions about the definition layers (Graphic 3.1). This is a direct consequence of the narrative structure of diaries and diaries commonly describe episodes of layer definition when noticing changing attributes. Technically, this isn't layer description, but the logic is based in the same principles.

> *When I removed U.7941 and came down upon a surface, we noticed that this surface had an edge and everything north of the edge was a 'surface' and everything to the south was fill (heterogeneous deposit with charcoal, small bone fragments and fallen bricks). So I went about removing the fill around the horns and the excavation confirmed our beliefs. I took out the fill to the back of the wall, which was still plastered. (SL/17/04).*

Many entries from B49's sequence describe various episodes of the removal of a big fill covering underlying features (Figure 3.3). These examples are particularly interesting because they describe the use of arbitrary levels for additional control when the extension of a layer seems to be not very secure and therefore this decision could have interpretative consequences.

> *Even though we are still in what we call "room fill", and the unit number should not change, theoretically, I have now become sufficiently worried about taking out almost 1m of fill in some areas to change the unit numbers in some places. These numbers still look very similar to what we have been excavating in the last two days, containing large pieces of mud brick, mortar, and various flecks of charcoal and plaster bits. (UK/07/04).*

Additionally, one testimony described the opposite strategy and instead of producing thinner slices, diggers lumped a group of layers into a context as in the example below. However, this testimony is less successful to explain the rationales of lumping. For example, if layers were visible but difficult to excavate or difficult to observe in principle. This information would have been useful to contextualize the reliability of context definition.

<u>Example 3.1</u>

> *It has been slightly unsatisfying since I didn't manage to take off the floor layer I had set out as the days project. Dan and I had some difficulties defining the lower boundary but in the end, we ended up lumping several layers and tomorrow it will hopefully all be gone.* (LF/44/08).

Furthermore, diaries describe episodes of uncertainty and fluidity related to the process of layer definition (Graphic 3.3). A recurrent example was uncertainty when defining layer boundaries as in the example below.

<u>Example 3.2</u>

> *We are down to features in several parts of the trench, but still have lots more fill to dig. It is hard to recognize some of it - as we get closer to the floor the fill is getting more compact and lighter in colour. While Ian and Shahina keep suggesting we mattock it out, the team is a little more cautious - we don't want to plow through any important features, and that is making us overly cautious.* (UK/04/04).

This example was crucial to revise some theoretical claims suggesting that ambiguities in layer boundaries always imply doubting the existence of deposits (Hodder 1999, 16). However, the former example illustrates that even if such ambiguity exists, this won't necessarily produce doubts regarding the existence of the fill and its underlying features. These examples were classified as persistent reflexivity or cases when uncertainty doesn't affect the credibility of interpretations. Nevertheless, there are some examples when uncertainty effectively affects the credibility of layer definition as in the following example which was defined as an instance of evidential reflexivity:

Example 3.3

> *So far, no in situ paint has been found, although at the bottom of the north wall (U. 7934) there is still collapsed wall plaster but I have been able to distinguish between wall fall and actual plaster wall (I think). […] it is a bit challenging to see as at the bottom levels, there is a plastered wall, a plastered platform and plastered wall fall all composed of identical materials, but the wall fall has irregular stratigraphy and looks out of place. (SL/11/04).*

A common characteristic of persistent and evidential reflexivity is that both types of doubt emerged due puzzling evidence. Yet, only one case affects the credibility of interpretative claims. These clearly contrasted with examples of trivial reflexivity describing episodes of uncertainty and fluidity produced by the changing observational conditions during the act of discovery as in the example below.

Example 3.4

> *I also investigated the cut in the east wall with 7945. As it turns out, it is not one cut but at least two. There is a late, almost rectangular cut visible in the plaster that goes through our side of the east wall, but not the abutting wall. Beneath this cut, there seems to be another change in the wall plaster which lines up with a disturbed area on the floor. While the date of the first cut is unclear, as only bone and obsidian was found in it), I suspect it is late Roman or Byzantine. The second cut might be Neolithic; I will therefore leave it until we come down to that layer.* (UK/21/04).

3.5.3 Finds

Diaries also describe the discovery and recovery of finds (Graphic 3.1). The level of detail in narration varies from case to case, but every testimony describes the type of find and its location. Likewise, entries also report when finds had been recovered by a different procedure: "*7938 and 7935 both yielded small clay animal figurines, which unfortunately came out of the sieve*" (UK/14/04). This is the only case study that includes reflexive information related to finds (Graphic 3.3). One Of the most distinctive examples of observational reflexivity was constituted by various episodes relating doubts and reconsiderations when defining a horncore. However, these entries also describe the progressive solution of those enigmas as remains are being cleared. This is one of the clearest examples that illustrates when ambiguity and fluidity are related to the changing observational conditions of excavation.

Example 3.5

The most exciting thing that we unearthed today was a plastered bull's horn possibly attached to the west wall. DGN found the horn core today while scraping what we previously thought was a post-hole or a post retrieval pit. So far, we believe that it is attached to the wall, although there is still about 5cm of mud covering the plaster. But there is a plaster line directly behind it, so it is not unlikely. (SL/02/04).

SLS excavated U. 7940, the remaining (artificial) platform around the horn core, and discovered that the horn was not attached to the wall. Oh well- it was an idea worth entertaining for a while. However, she did find more of the horn exposed at a slightly different level and will continue to expose it tomorrow. (SL/13/04).

More progress today in the space, and many more finds. The horn core 7920[3] *was exposed to be much larger than originally thought. (UK/15/04)*

Diaries showed that layer and finds detection is one of the aspects where interactivity with laboratory specialists had a positive effect for defining layers and finds due the role of additional instruments, occasionally in situations of doubt. For that reason, being examples of evidential reflexivity, although the impact of specialist onsite was elementary as it only helped diggers to observe better, not to interpret more.

Example 3.6

On the western wall the painting appears to be very early in the sequence, a micro-morphology sample taken from the wall will allow the number of layers of plaster to be counted. Initial examination with the naked eye suggests and initial thick foundation layer was put on the wall, the first few layers of plaster contain the painted designs, these are then sealed by a second thick make up layer and then a series of typically thin re-plastering events. (DE/30/06).

Finally, some entries describe episodes of uncertainty about find’s attributes addressed with specialists. Yet, none of these examples was reconsidered in subsequent entries. Mainly, because

[3] Some entries referred to the horncore as the ‘attached horncore’ whereas other as U.7920. The connection between those such testimonies became evident with an entry that included both: “We did get a number of good results: the horn U.7920 turned out not to be attached to the east wall” (UK/14/04).

such doubts didn't seem to have interpretative relevance, thereby being classified as persistent reflexivity.

In U. 7926 we came down upon another horn core (U. 7928). There was some question as to whether or not it was plastered. Upon excavation, we all thought it was plastered. Ian and Shahina both thought it was salt and not plaster. Wendy came around on the priority tour and took a sample- the result is still pending. (SL/08/04).

3.5.4 Other Low-level claims

Occasionally, entries mention similarities and differences among layers and finds (Low-level contrasts). For instance, the size and preservation of two horncores. However, the most interesting examples use these observations to guide interpretations, for example problematic layer distinctions (Example 3.3). Diaries also includes a few examples of Expected Absences or claims describing situations when an expected finds or layer attribute wasn't observed, "*No grave good were observed in the burial*" (DE/29/06). Occasionally, these claims might have an important role in reflexive testimonies to evaluate interpretive possibilities (Example 3.9). Finally, there are cases of Indicative Absence or claims describing situations when an absent sign is considered to be evidence of something else as in the example below. The proportion of these claims was very low, and all of them were quantified in one category as 'other low level claims' (Graphic 3.1). A common characteristic of these claims is that they can only emerge in the light of some form of background knowledge which enables comparing things and have expectations.

"Two floors were removed from platform F1666, sealed the basin. These in turn sealed a small pit cut [13648], possibly representing an emptied-out cache, the fill of the pit contained no significant finds" (DE/29/06).

3.5.5 Stratigraphic information

Stratigraphic information was not very frequent in diaries (Graphic 3.1). In general, layer contacts is the most frequent type of stratigraphic information because these claims usually structure the narrative of layer discovery as in the example below.

The two burials are cutting platform surface and central floor 16636 which was partly taken off (the rest will be removed when the burials are out). Floor 16636 seals the larger central floor 16646 which was today's project and which will continue tomorrow. (LF/44/08)

Very occasionally, entries describe stratigraphic relations. In such cases, entries normally indicate whether a layer was earlier/later in relation to another, hence these examples always report relatively simple sequence interpretations.

> *The plaster was removed from sections of the northern and western walls of building 49, directly above the NW platform F1651. Several layers of painted plaster were encountered. On the northern later layer of paint appeared to consist of a solid red colour. Earlier layers 13669 contained geometric designs in red and* black paint. (DE/30/06).

When diaries include reflexive situations related to stratigraphy, these generally describe situations of uncertainty and fluidity in layer contacts during the process of discovery. However, as observational circumstances changed doubts were solved. Hence, these examples were classified as observational reflexivity (Graphic 3.3).

> *The one relationship that we have not yet been able to determine is its association or connection with the southwest partition wall. This connection is not as clear because there is a plastered brick/surface jutting out and we are not clear what this brick is associated with. Also, the arm to the other basin (U.7950) is intersecting F.1652. Again, we have not been able to determine which of these basins came first or if they are contemporary; they might be contemporary because they both appear to have been built on the same truncated surface but the connection to the partition wall and the south wall are still unexcavated and unclear. I excavated a little more of the plaster line of the arms of the basin U.7950 and was able to establish that the plaster connects with the plastered surface (U.7944) and it seems to connect with the line running across the space between the two surviving bits of what we are calling the partition wall. (SL/26/04).*

Only a small portion of reflexive testimonies describe interpretive doubts in layer contacts produced by puzzling evidence. For example, the testimony below describes an ambiguous relation between 16550 and F4017 due fragmentary evidence. Although, the testimony fails to explain the supporting or contradictory evidence for the stratigraphic relation. Therefore, being classified as an example of evidential reflexivity but being a problematic testimony for having inadequate standards to report information (Graphic 3.3).

> *F4016 was altered by the addition of clay silt make up (16650) to the N end, possibly associated with the construction of F4017. The extent, nature and relationship of (16650) were unclear as a large animal burrow to the north heavily truncated it. (DE/46/08)*

3.5.6 Assemblages features and structural elements

Assembles, features and structures formed the main internal elements of B49 (Graphic 3.1). And diaries include testimonies describing the discovery of these types of vestiges. Assemblage claims usually describe the discovery of human burial with grave goods or clusters of finds spatially related. For example, the following testimony describes the discovery of a figurine's assemblage:

<u>Example 3.7</u>

> *Today was quite a day for finds! I started excavating U. 7957 in the south west corner, behind the 'partition wall'. Initially, I wanted to follow the plastered line that had been appearing on the surface for a wall and see how it connected with U. 7954. Within minutes of starting, I had popped up several animal figurines and put them to the side. I had not really considered them anything special until I had mentioned it to Emma, who showed great excitement for them. There were four in the area just west of the plaster line, inside the 'structure' and I clustered them as U.7958. One of the figurines had popped up several days ago and had just been sitting there waiting for us to excavate the unit. The last figurine was found about 10 cm from the cluster. UK had found 2 or 3 similar figurines in the units above U. 7957, (7935 x1, 7938 x1,x4). Who knew that finding a figurine cache would be such a pain? (SL/25/04).*

Structural information was among the commonest. This Includes short descriptions of structural features like in the following example: "*Platform F1651 is still a large and dominant feature in the room, and traces of red paint on the eastern vertical face indicate it was painted red*" (DE/43/08). However, the most common account are descriptions of episodes in the discovery of mud brick-walls, plaster layers, and makeup layers forming part of different types of masonry features (ovens, hearth and platforms). Commonly, these episodes describe doubts and ambiguities related to the discovery of structures, but many of these testimonies also include predictive conjectures when fieldworkers make an initial interpretation about the type of structure being excavated. One of the characteristic examples was a series of related entries describing the discovery of an oven (Example 3.8). These examples usually report the earliest stages of definition or until the exposition of remains enabled a secure interpretation of its type. Hence, having a full testimony of the discovery of structures was unnecessary. Once more, these examples were classified as observational reflexivity because doubts and ambiguities are produced by observational conditions (Graphic 3.3).

<u>Example 3.8</u>

> *We took out a great deal of volume today and hopefully tomorrow we will be able to more clearly outline the area around the other horn core (U. 7920). I expect to find an oven, or perhaps a hearth, in this area but we have not seen anything as of yet. There is a very slight hint of a circular structure, but it was the end of the day, so we'll examine it a little more closely tomorrow. (SL/08/04).*

> *As I came down to the current level of excavation in the south east corner, I noticed there were several burnt mud bricks, similar to what I would expect an oven to look like. (SL/13/04).*

> *The conservators took out the last two horn cores, 7920 1 and 4. Mira and Ruth gave input on the southern area of the space, which has been extremely confusing with its various layers of plastered surfaces, feature walls, and burnt material. The current favourite theory is that we actually are seeing TWO hearths/ovens, not just one. One is under 7943 - we can just see the top bricks and the arms coming out. This one may not be attached to the south wall. The other one we can see so far only in the presence of heavily burnt gunk in front of the bench/platform in the southeast corner (UK/19/04).*

Less often, diaries include examples of retrodictive interpretations after exposing structures which clearly contrasted which predictive situations based in a limited visibility of structures. Occasionally, retrodictive conjectures were closely related with reflexive testimonies describing interpretative reconsiderations like in the example below that discusses whether a feature might be a posthole or a bin. The example below is the most paradigmatic case of evidential reflexivity in which a fieldworker develops an argumentative discussion of evidence assessing supporting and contradictory evidence for interpretative possibilities. This example was crucial to notice more clearly alternative ways to achieve a reflexive recording, one where fieldworkers describes doubts and reconsiderations produced by the act of discovery and a second way when testimonies describe doubts that persist after the act of discovery because evidence is puzzling and compromises the credibility of interpretations.

<u>Example 3.9</u>

> *What was originally recorded as post retrieval pit F1495 in the NW corner of building 49 was found, upon further investigation, to be slightly unusual. Vertical wall plaster 13698, 13676 and 13675 obviously enclosed the feature however the inside of the feature, where a post should have stood, was also plastered. This presents two possibilities, firstly an earlier basin was reused as the base for a plastered upright post, however the post was not dug in and*

> *left no scar on the back wall. It is more likely the feature was in fact some kind of tall thin bin. Possible uses for such a feature include storage or the feature may have held a lamp or candle. The plaster behind the feature is covered in a thin black greasy residue, this is currently unexcavated but may allow the function of the feature to be identified next season. (DE/30/06).*

Additionally, there is an example of interactivity during the interpretation of an ambiguous structure. This example is important because it shows the effects of information exchange in priority tours, but it also shows the importance of previous experience in the interpretation of remains.

<u>Example 3.10</u>

> *Dragona, who came by as part of the priority tour, made a useful suggestion for the irregular platform-type feature, backed by bricks, that I had tried to define with 7942 yesterday: apparently, it looks a lot like a ladder/stair support that BACH had found in their space a few years back. (UK/18/04)*

Occasionally, diaries report cases of uncertainty of structural attributes when evidence was fragmentary. For instance, the testimony below describes a doubt about the extension of a platform, yet the type of structure is secure. Hence, these examples represent instances of persistent reflexivity because such doubts do not compromise the credibility of interpretations. Furthermore, unlike cases of evidential reflexivity, testimonies of persistent reflexivity do not seem to require further discussion.

<u>Example 3.11</u>

> *The platform associated with F4010 was heavily truncated prior to being remodelled and incorporated into platform F4008. The original form of this platform was therefore unclear. (DE/46/08)*

Burials were the commonest type of non-structural feature, but diaries do not describe the discovery of these vestiges very often. Besides, testimonies of evidential reflexivity are uncommon for assemblages and features, there is only one testimony describing uncertainty if a find belong to an assemblage (see below). Yet, this testimony fails to explain the cause of uncertainty, for example if the artefact was located between two layers or if it seemed intrusive. Likewise, it fails to discuss supportive and contradictory evidence. Otherwise, this information would have been very useful to contextualize the testimony and the credibility of the interpretation. Therefore, this example represents another instance of an inadequate recording standard for a testimony of evidential reflexivity (Graphic 3.3).

<u>Example 3.12</u>

> *I started excavating platform F.1651, located in the northwest corner of the building, which had a very large round impression in the centre of the plaster, layer 13668. I took off a series of red make-up, dirty surfaces, and white plaster, revealing the top of burial cut [14437], starting f. 4000. After excavating the burial fill (14429), skeletons 14441 (a young woman) and 14440 (an infant) were revealed by Lori Hagar. Along with the young female skeleton were a number of ground stone beads closely associated with the neck of the skeleton. There was a small greenstone axe in the fill, possibly associated with the infant 14440. (CM/37/08).*

3.5.7 Formation processes

Claims about formation processes was another component of reflexive diaries (Graphic 3.1). These data usually describe use, modification and destruction events in masonry remains. In that sense, there wasn't much diversity because the interpretation of depositional process was largely constrained to structures. This characteristic was indirectly acknowledged in some entries: "*internal wall F1654, like so much of space 100, consists of a number of rebuilds and modifications*" (DE/30/06). Generally, formation process claims describe relatively simple retrodictive inferences based in stratigraphic data.

> *I have been working on some of the wall paintings in the NW area of the building as well as excavating some smaller architectural elements. An internal division or small bench F4016 consisted of an earlier feature, possibly a platform in the SW corner of the building, which had been cut back to create a sunken storage area. F4016 was constructed of relatively thick layers of plaster and brown clay silt make up (16674) and (16672), which represent repeated construction events on an earlier truncated feature. These were overlain by brown make up (16671) deposited against the western side of the feature. This provided the shape of the feature and was sealed by clay plaster (16669). F4016 was altered by the addition of clay silt make up (16650) to the N end, possibly associated with the construction of F4017. (DE/46/08).*

But occasionally, depositional processes were interpreted from subtler signs: "*The building was obviously occupied for a considerable period of time and the patchy worn floors across the centre of the building suggest heavy and regular use*". (DE/30/06). Including a few examples of depositional investigation for non-structural remains:

In the northeast corner, I found that the burial cut (7917) had extended through our plastered wall. This is interesting because the skeleton itself did not extend this far, leading me to think that perhaps the grave was originally cut for an adult but then used for a child. (SL/13/04).

In most cases, depositional data include an explanatory conjecture linking evidence with a depositional interpretation, except for a few cases of unjustified interpretations. Like in the example below that describes the primary character of a layer without providing favourable evidence. Equally interesting is that similar cases usually correspond to soil layers.

A white plaster floor (12693) was removed from platform F1656 to reveal an ashy deposit (12698) containing fishbone, a piece of worked bone (12699-X1) and a cluster of obsidian flakes (13600), which represent knapping activity. (DE/28/06)

Formation processes is the clearest area where one can appreciate the benefits of interactivity between fieldworkers and specialist in interpretative work. Likewise, these examples show very clearly the development of a feedback loop in which data from the laboratories contributes to reinforce or revise a preliminary field interpretation. Besides, unlike sampling information, which is generally unclear, diaries are more effective to explain the reflexive process when fieldworkers receive laboratory data. The most interesting examples describe situations when a preliminary interpretation demands revision in the light of new evidence (Example 3.13). Hence, these are the clearest examples of evidential reflexivity in cases of fluidity. In fact, this was the most frequent form of evidential reflexivity (Graphic 3.3).

<u>Example 3.13</u>

While defining the surface of the platform, I popped up a piece of what looks like blue paint. The colour was found in a lamination layer. Ina had a look at the "blue" paint and took a thin sample of it (Sample #5, unit 7913). I removed the remainder of the block and sent it to conservation at the end of the day. After lunch, Wendy Matthews came and had a look at all the areas with lamination. I had originally thought that they were just thick layers of painted plaster that had fallen down, mainly because of the way in which the pieces were haphazardly situated. After several of the pieces had come up, I was concerned that they may actually have been something else. Wendy came up and confirmed that these laminated pieces were indeed fallen plaster and asked that we saved some of the samples. She also mentioned that the "blue" colour may actually be smoke 'damage'. (SL/06/04)

3.5.8 Building and spaces

From the start of the exploration, it was evident that B49 was a building, hence this was not a recurrent topic in diaries except as general reference: "*First days of work in building 49 in the 4040 area*" (LF/32/08). Likewise, from the earliest stage, it was obvious that B49 contained an internal area defined as Space 100 (Figure 3.4). As the excavation continued, and internal walls were discovered, diggers began to ponder about the existence of an additional room. The fragments below represent a selection of related testimonies describing the discovery of an internal wall (F1654) and the definition of an additional room (Figure 3.5).

<u>Example 3.14</u>

The cleaning of the walls also revealed that the wall south of the "bin" was plastered with much thicker plaster than the west wall. Discovering this, DGN interpreted this to mean that the bin wall actually functioned as the inside wall of the living space, while the west wall was the inside of a "storage area" and only got replastered infrequently. (UK/12/04).

To replace our missing students, we borrowed Val from Berkeley and he worked inside the bin area (U. 7921) and it looks like he might have a plastered surface at the bottom of it. We'll explore it further tomorrow. A large amount of animal bones has come out of this unit but the soil matrix still looks like fill and the outside walls are plastered. No real thoughts yet as to what this structure thing might be. (SL/20/04).

VS continued 7921 today - it is looking more and more like a small side room, although a large amount of fallen brick and a plethora of rodent burrows obscure its form. What we thought was a bin floor in 7921 yesterday today turned out to not be one - it does not reach the west wall as far as we can tell. Perhaps it was a part of the dividing wall that then fell over? (UK/21/04).

Further clearing revealed an additional fragment of the dividing wall (F1659) and this allowed to observe more directly the existence of an additional space. At this stage, diggers didn't include an extensive narrative of discovery and they simply described the spatial organization of the building in the following terms:

We have removed an internal partition wall F1654 and F1659, which divided space 100 from a small side room, recorded as space 334. The central area of the internal wall was heavily truncated by animal burrows. The northern part of the wall was recorded as F1654. (DE/35/08).

Most of the uncertainty and fluidity claims related to space definition were classified as observational reflexivity as this was a direct consequence of the discovery of internal walls (Graphic 3.3). However, some of these claims were classified as evidential reflexivity because they describe interpretative doubts produced by the fragmentariness of remains. Yet, when both wall fragments were exposed, the spatial organization of the building was relatively evident (Figure 3.3).

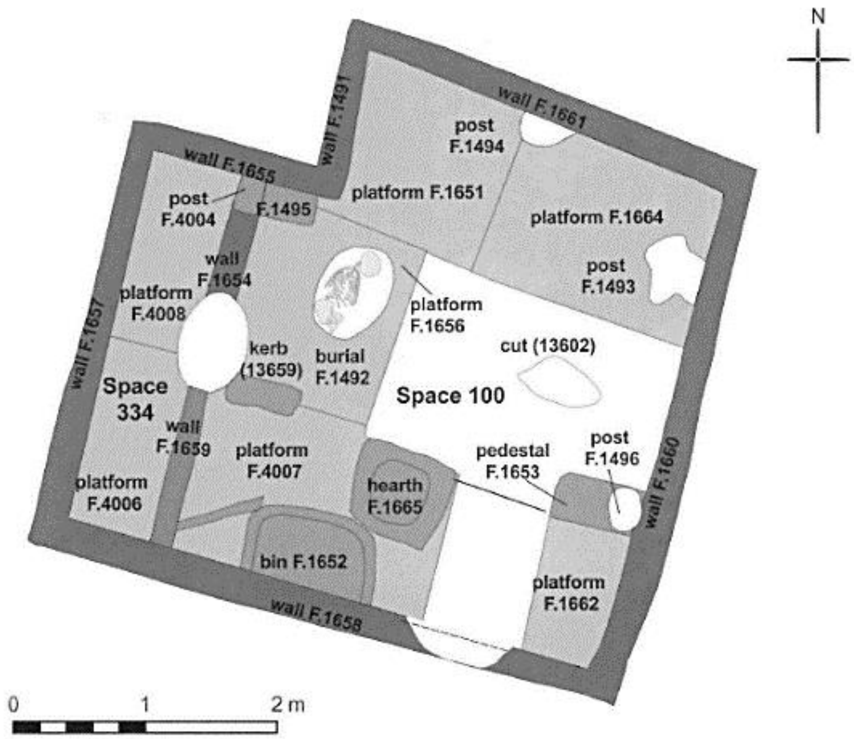


Figure 3.5. Building 49 (Copyright Catalhöyük Research Project).

Diaries also include some descriptions of activity areas in the main room. These interpretations are based in the ashy-ness/cleanness of features (Middle level contrast). Besides, some entries remark the regularity of this pattern across the site, indicating similarities between B49 and other buildings. Again, these testimonies illustrate the importance of previous knowledge gained by past experience.

> *The typical pattern of dirtier floors in the south of the building, associated with hearths and ovens, and cleaner white floors in the north appears to replicate here. The central floor area of the building appears to represent more mixed use with both clean and dirty floors. (DE/30/06)*

Platform F4007 consist of an upper and lower level. On the upper level - the visible platform area - the floors and makeup layers are less ashy than in direct vicinity of the oven, but clearly darker and more ashy than the plaster floors in the central area of the building space 335. (LF/38/08).

3.5.9 Constructive phases

Diaries include a small proportion of interpretative data about constructive phases (Graphic 3.1). Most of these claims are based in the evidence of a repetitive constructive behaviour and the high preservation of evidence. Moreover, phase claims move away from the narrative format and turn into synthesis of multiple types of empirical observations, besides, the most complex examples make use of previous observations in the area.

It would seem that there are two rough phases: floors that have been removed BEFORE[4] the larger plaster floor (14492) from the central area joins up and run over the platform Units removed from upper area: 14423, 14464, 14466, 14478, 14491. Units removed from lower area: 14467, 14468, 14469, 14472, 144474 and floors removed AFTER the sealing of floor (14492). Units removed from upper area: 14495, 14496, 14497, 16604 units removed from lower area: 14498, 16603. (LF/38/08).

Clearly the engaged pillar and the adjoining north wall was repeatedly painted with red and black geometric designs. The earliest of these recorded to date, (16605), consisted of a pattern of black diamonds and red decoration. [...]*– it certainly appears that there is at least one earlier phase of painting. Layer (16605) was sealed by white plaster (16608), representing a number of re-plastering events. This was sealed by further layers of red and black painted plaster (16607)* [...] *During the 2006 excavation season a series of painted plasters were exposed and recorded on the north and northwest walls, around the northern platforms F1651 and F1654. These paintings are all earlier than the current building phase and appear to have been painted and plastered over very early in the building sequence. Certainly, during the current phase, which includes several burials cut into the northern platforms, the northern walls F1491 and F 1661 were plastered white with no painted decoration (DE/39/08).*

In the earliest season, diggers reported some interpretive doubts produced by observational conditions during stratigraphic exploration as indicated in the fragment below. Hence, this example was classified as instance of observational reflexivity (Graphic 3.3).

---

[4] Emphasis in the original

> *The presence of yet more features turned out to be more confusing still. We seem to be down, now, on at least 5 different use levels just in the features, not counting the washes on the bin sides and the floors. (UK/10/04).*

In the last season, a digger reported an interpretative reconsideration. This was the best example of a testimony of evidential reflexivity for phase interpretation. Yet, this claim is very limited as it fails to explain supporting evidence for such interpretative change. Hence this represents another clear example of a recording problem in evidential testimonies that lacks a more detailed discussion of evidence. Overall, phase claims represent one of the clearest examples of the best potential of diaries for recording complex data. However, this requires being able to synthesize and explain a los evidence.

> *In this new phase an oven, F4003, recessed in a niche in the SE corner of the building appears to have been in use. Originally, I thought this oven was in use in the previous phase however it now appears that hearth F1665 was in use until now. Units (14470) and (14499) appear to represent an earliest phase of the hearth that was missed in the 2006 season. Therefore, several of the 'dirty' floors associated with oven F4003 may have been in fact associated with hearth F1665 (DE/43/08).*

3.5.10 High-level claims

High-level claims are the less frequent class of information. These include functional, agency and culture/date claims (Graphic 3.1). Culture/Date claims include a couple of examples that describe if a remain was byzantine, roman or Neolithic. These data are more common in the early stages of exploration when diggers removed a few roman and byzantine features above the Neolithic building. In general these interpretations lacked a justification, occasionally some of these testimonies indicated doubts about the cultural affiliation of remains (Example 3.4) which in general never included a discussion of evidence. This was common aspect with additional projects. However, reflexive claims were considered as examples of an inadequate recording given the characteristics and goals of reflexive method (Graphic 3.3). Particularly, SCR and the feature system consider that this type of claims are the domain of post-excavation whereas the reflexive method aims to erase that division.

Agency claims only include two examples of interpreting evidence in terms of human behaviour. In one entry, a fieldworker interprets a midden deposit as a possible cleaning action: "*This soft midden like deposit may represent a house clearance event*, ***with the house cleared of older*** *materials, hearths emptied etc*". (DE/36/08). However, this a problematic statement because empirical support is not very clear. Besides, this is the only example of an interpretative doubt for an agency claim and therefore was considered an instance of inadequate recording (Graphic 3.3). A more interesting example is when the same fieldworker makes a more general statement that reinterprets architectural changes in terms of social and emotional behaviour.

> *The stratigraphy of building 49 is particularly complex and represents the constant alteration and modification of the space with what appears to be puzzling frequency. As well as the addition of new floors and features there is also a constant cutting back and re-modelling of earlier features, as if the occupants are never quite happy with their living environment. (DE/30/06).*

Finally, functional claims are not very frequent, although they are the most frequent type of high-level claim (Graphic 3.1). Generally, functional claims refer to specific parts of the building such structures and spaces (see fragment below), but it's not uncommon that functional interpretations are unjustified. The most interesting case of this type of failure the fewest testimonies describing the functional status of the building: "*Building 49 is a small house, with plaster and the traditional features that characterize Catalhöyük*" (CM/37/08). This case is particularly important because the domestic status of B49 is the most important interpretation of the building. Yet, it is never explained or discussed. Equally, some of these examples describe functional doubts but testimonies frequently lack a discussion of evidence as in the following example (Graphic 3.3).

> *Basin F4010 was located on a platform in the NW corner and probably associated with an area of charcoal and dirty floors. This area appears to represent an activity area, possibly associated with food preparation (??). (DE/46/08).*

3.6 Reconsidering reflexivity

So far, one of the main results of the analysis of B49's diaries is the classification of reflexive claims into three categories: evidential, persistent and observational reflexivity (Graphic 3.3). These categories have been useful to observe different types of interpretative doubts and reconsiderations. Previous examples have been useful to sketch these categories but now is time to provide preliminary definitions. <u>Observational reflexivity</u> includes cases of uncertainty, fluidity and possibility produced by observational circumstances, specifically the changing threshold of what is visible and invisible during the act of discovery. Generally, these testimonies show a repetitive pattern which often begins with an interpretative doubt or a preliminary hypothesis about a thing being observed and partially exposed. As excavation continues, there might be some reconsiderations of previous observations. However, interpretive doubts are generally solved when remains have been exposed. For this reason, the narrative format perfectly fits to describe a sequence of excavation episodes that altogether provide a testimony of the act of discovery, likewise predictive conjectures make perfect sense in this type of testimony, especially when the visibility of things is restrained. Overall, this form of reflexivity is related to observational situations when defining layers, assemblages, features and spaces. So far, the most paradigmatic examples are the extensive chronicles of discovery of a horncore (Example 3.5), an oven (Example 3.8) and the side room (Example 3.14).

Evidential reflexivity includes cases of interpretive doubts produced by puzzling evidence. In the case of evidential reflexivity, observational circumstances are irrelevant because these claims are made when a vestige has been exposed. The most paradigmatic examples include a case of uncertainty in the number of layers being observed (Example 3.3) and when the type and function of a feature was ambiguous (Example 3.9). Additionally, cases of evidential fluidity emerge when a retrodictive conjecture is revised in the light of new evidence. So far, the most paradigmatic example describes the revision of formation processes (Example 3.13). The best examples of evidential reflexivity have an argumentative format in which fieldworkers discuss supporting and contradictory evidence to assess interpretative possibilities or to explain interpretive reconsiderations. In short, testimonies of evidential reflexivity discussing puzzling evidence or new data when that compromises the credibility of an interpretation.

Overall, observational and evidential aspects define a reflexive cycle constituted by states of doubt and belief. Generally, doubt is triggered by new observations that demand explanation; then when observations are explained the state of doubt is soothed. However, the cycle might reactivate with additional information (Aliseda 2005, 368). In general, this cycle is well accounted by hermeneutic theory (Hodder 1999, 33) and the theory of the act of discovery (Edgeworth 2012). Yet, observational and evidential reflexivity differentiate if the state of doubt is soothed or not. Specifically, observational reflexivity represents cases when interpretive doubts dissipate as visibility improves. On the contrary, doubt persists in evidential situations because evidence is puzzling (uncertainty) or doubt reactivates with new data (fluidity), hence affecting the credibility of data. Finally, Persistent reflexivity, is another type interpretive doubt produced by puzzling evidence but unlike cases of evidential reflexivity, persistent doubts don't compromise the credibility of data. For example, cases of ambiguous layers without producing controversy in layer definition (Example 3.2) or cases of fragmentary evidence without compromising the confidence of an observation (Example 3.11).

These reflexive categories are inspired in previous theoretical elements. For instance, the distinction between predictive and retrodictive conjectures as when tracking an animal or reconstructing an event, respectively (Ginzburg 1989, 102). Thereby, there is a resemblance in the process of tracking an animal and a cut from its visible clues (Edgeworth 2012, 79). However, when one finds the animal or discovers the cut, one no longer has evidence of such things, one sees them directly (Austin 1962, 115). In contrasts, cases of evidential reflexivity generally discuss aspects that in principle are unobservable and evidence is puzzling therefore affecting the credibility of interpretations. For example, Natalie Zemon Davis says that documents can provide sound evidence about some past phenomena but be more fragmentary about others which will be subject to various interpretative possibilities that the historian will relate as the "perhapses" and the "may-have-beens" when evidence is inadequate or perplexing (Zemon-Davis 1983, 53).

Finally, the distinction between observational and evidential reflexivity not only differentiates two epistemic situations. It also serves to differentiate between relevant and trivial information in reflexive diaries. In general, Hodder and his colleagues maintain that reflexive diaries relevant to give a testimony of the act of discovery which will support the credibility of primary data (Hodder 2000a, 9, 2005, 651) or as Alison Mickel asserts: "*The entry grants access to transitory moments that are traditionally black-boxed in technical reports and proforma*" (Mickel 2015, 303). Furthermore, in a recent study, Mickel examined compared the content of diaries to explain their relevance (Mickel 2015). Specifically, she classified the content of diaries in two classes: distinctive and redundant information. Redundant information included various types of claims found in context sheets and diaries from Catalhöyük alike, this includes descriptions of layers, features and stratigraphic data. On the contrary, distinctive information included information exclusively found in diaries such as chronicles of excavation, discussions of finds and methods; and hypothesis (Mickel 2015, 302). In broad terms, redundant information corresponds to what this study defines as interpretative claims. Whereas, distinctive information fits with reflexive and non-interpretative content. Likewise, both studies present similar quantitative data as distinctive information is less frequent than redundant information. Then, there is a nice consistency in the description of diaries' content.

One of the main arguments of reflexive archaeologists to defend the relevance of diaries is their distinctive content which allegedly contributes for the credibility of data. Specifically, Mickel stresses the role of narratives of discovery: *"Only the diary preserved these interpretative steps, permitting those not present at excavation to enter into and evaluate each stage of the reasoning process*" (Mickel 2015, 303). However, this argument is not very compelling because it fails to explain the way in which such narratives reinforce traditional testimonies. For instance, the way in which a description of the discovery of a wall reinforces a typical testimony of having observed a wall given by a description, a drawing or a photograph (Example 3.14). Hence, a) the evidential function of narratives relating the changing impressions during excavation is unclear, b) especially when the characteristics of remains become evident after being exposed. This is particularly recurrent at Catalhöyük where the preservation of remains is very high. Finally, if one considers that reflexive diaries are primarily textual records with no sketches. Then, narratives of discovery seem redundant because a sketch or a photo are more economic and more efficient resources for this task. In other words, there are not a good reason to believe that a reflexive narrative contributes to stablish the credibility of having observed something, in this case a building, its internal structures and spaces (Figure 3.3). The epistemic weight of textual over visual testimonies is an aspect that will be examined in more detail later (see section 6.6).

In contrast, the most valuable data in reflexive diaries are the testimonies of evidential reflexivity that describe the reasoning process of fieldworkers when marginal and fragmentary evidence produces puzzlement and therefore compromising the credibility of data which require a more detailed discussion of evidence and interpretations. Similarly, another important element of diaries are the testimonies of interactivity among fieldworkers and specialists. In this case, diaries clearly demonstrate the advantages of having immediate feedback onsite, which improves the epistemic circumstances of fieldworkers with a wider background knowledge to revise previous ideas. Overall, the main problem with Mickel is failing to notice the importance of discussions of evidence over chronicles of discovery. Instead, she presumes that having more information implies having better testimonies. For these reasons, reflexive recording as defined by Hodder and his intellectual progeny is simplistic.

Finally, there is a minor issue regarding redundancy of data in excavation diaries and context sheets from Catalhöyük. Mickel assumes that the most distinctive information of diaries is absent in context sheets, namely narratives and discussions of evidence (Mickel 2015, 303). However, after having noticed the greater epistemic value of discussions of evidence, I checked a few context sheets related to one of the most important reflexive testimonies (Example 3.9). There I found, another version of the reflexive discussion elaborated in diaries, specifically in context sheet [136664], one of the fills of feature F1495 (see fragment below). Although, no systematic analysis of context sheets from B49 was made, this limited evidence raises some doubts about the idea that reflexive discussions are absent form standardized formats. If this is the case, one might another reason to believe in the irrelevance of Catalhöyük excavation diaries.

> *Fill of post retrieval pit, includes collapsed plaster from the engaged pillar. The post pit is not very deep and appears to just sit on the plaster floor rather than being dug into it. Several small fragments of animal bone and obsidian were found, and a shell was found in the SE corner but was probably disturbed by animal burrowing activity. Further excavation by DE revealed that this feature is unlikely to be an engaged pillar and is more likely to be a tall narrow bin. The feature was repeatedly plastered on both the inside and outside and traces of red paint were also found. The base was also plastered on more than one occasion and there is no evidence of a post-pit or scarring on the wall*[5].

The reflexive investigation of B49 and its epistemic context

In the previous section a general assessment of methodological aspects has been presented. However, the internal analysis has allowed to observe some problems to match the recording process and the epistemic context of the investigation. As explained before, B49 was a very well-preserved building clearly identified in the early stages of its excavation (Figure 3.4). Besides, B49 was interpreted as a domestic structure at the same time: "*We cannot tell with any certainty to which Neolithic level this house -Space 100- belongs*" (Krotscheck 2003, 23). This assumption was based in previous exploration by the CRP which indicated a domestic function of buildings (Hodder and Farid 2014, 3). This idea contradicted previous results by James Mellaart claiming that buildings included a mix of houses and shrines (Mellaart 1967, 3). Nevertheless, B49's diaries never discuss the domestic character of the building (see section 3.5.10). And this is an

[5] http://db.catalhoyuk.com/database/catal/UnitSheet.asp?num=13664

important problem because although it was evident that B49 was a building, it was not evident that such building was a house. Then by uncritically embracing the premise of the domestic character of B49, one of the main goals of reflexivity was left aside, namely the examination of archaeological assumptions (Hodder 2000a, 9).

In other words, the obvious presence of a B49 makes irrelevant a narrative of discovery that largely reiterates what is already known. Moreover, even if a photo was a sound resource to demonstrate the existence of B49, it was insufficient to prove its domestic nature, and this another important reason for prioritizing reflexive discussions. B49 diaries include some glimpses of reflexive discussions addressing functional ambiguity for specific features. However, a more systematic strategy would have produced a more valuable record. Specifically, the CRP had favourable conditions to achieve this task given the available background knowledge, the character of the site and the tools at hand. Altogether, this demonstrates that another important problem with reflexive diaries is failing to consider the epistemic context of an investigation in the production of site records.

Methodological reflexivity

Additionally, there was another type of reflexive information in B49 diaries, namely methodological claims (Graphic 3.1). The proportion of methodological claims was very similar to other reflexive claims, but methodological information only includes two extensive testimonies that address recording procedures during the first season. One was the only entry recorded by a trainee that described some problematic aspects when filling up sheets. However, it's not very clear from this testimony whether these are legitimate methodological problems or whether such 'confusions' are the consequence of unfamiliarity with recording rules, specially how to fill and update excavation sheets.

> *What is confusing to me is that unexcavated (what I will refer to as) 'definitional' units can be included in "feature units" for an excavated feature. It is important to discuss these 'definitional' units because they may support or bound excavated units that comprised the excavated feature, and they are visible at the end and can be useful for interpretation of what has been excavated* […] *Also, the interpretation of an unexcavated 'definitional' unit may change with its later excavation, possibly invalidating interpretations made for a feature it does not make up. These issues are not problematic if everyone is aware of them, but I think adding checkboxes on the feature sheet*

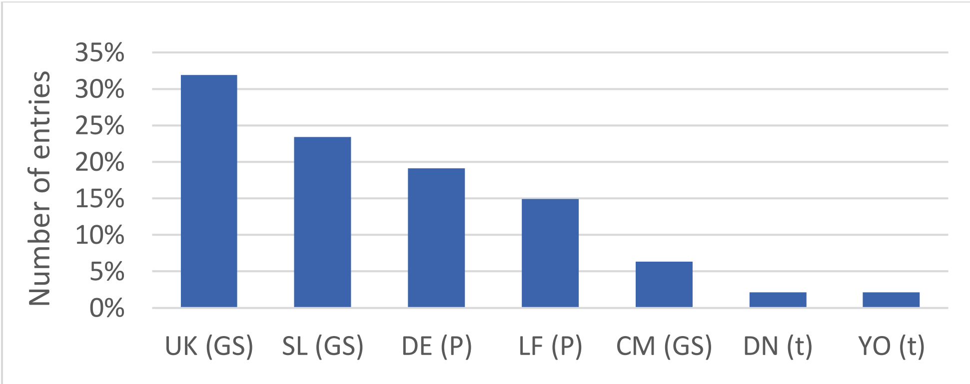
for excavated/unexcavated and compositional/definitional in the "feature units" section would be useful. (DN/27/04).

Another entry makes similar observations about how diary's content didn't reflect important discussions onsite (see below). Possibly, this testimony is a very rudimentary version of the more detailed analysis presented in this chapter, but evidently it doesn't provide an accurate assessment of field practice. Independently of this, excavation diaries effectively provide a space for fieldworkers to reflect upon their practice. Nevertheless, these voices speak in the dessert because the architects of reflexive method are more interested in technology (Berggren et al. 2015). Hence, fieldworkers experience doesn't seem to have any incidence in the revision of methodological principles. Arguably, a closer communication between fieldworkers and directors, would have facilitated noticing the importance of discussion over chronicles.

However, one thing I did notice today was that I wish I had been more diligent and detailed with my diary entries. We have had so many conversations in the space concerning the relationships between units and features and how thing relate to each other in phase. Only speaking for myself, I am not convinced that my previous entries did any of those 'real time' conversations any justice. I suppose that is the point of a 'reflexive methodology'- to live and learn (or trowel and error) (SL/26/04).

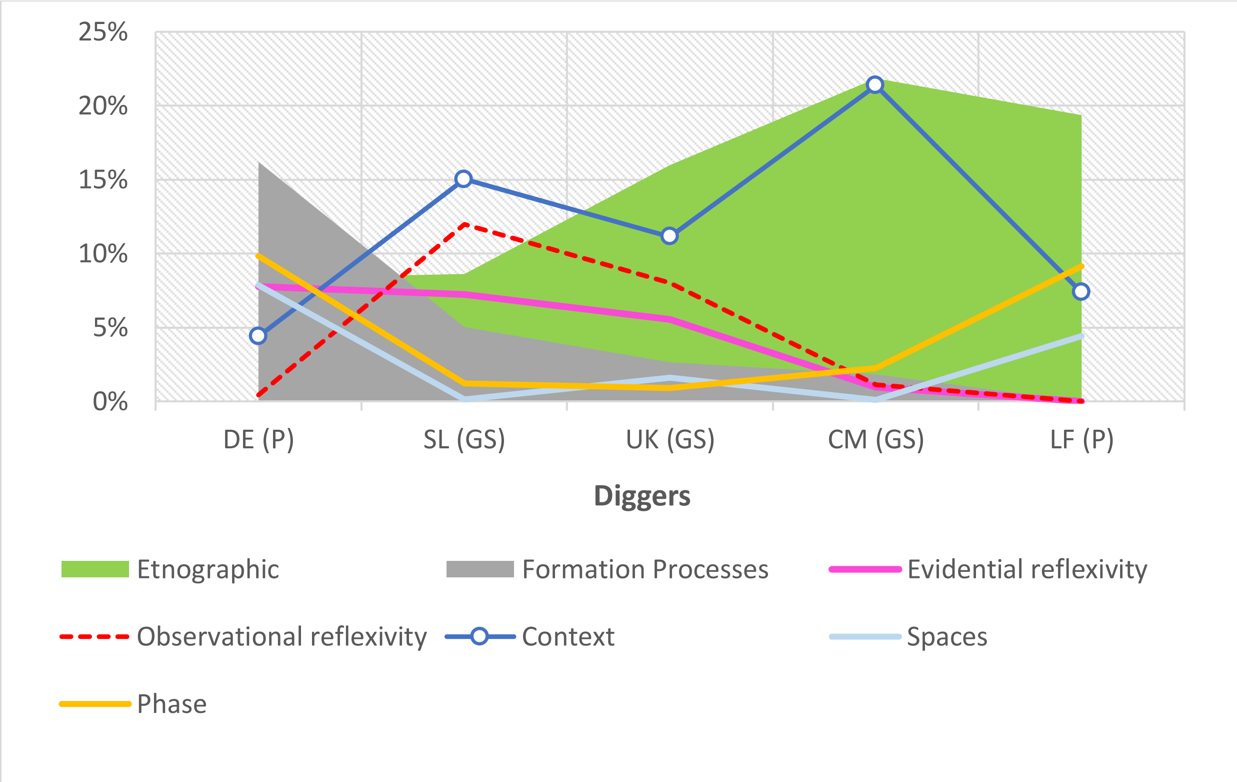
3.7 The performance of fieldworkers

B49's diaries include testimonies of seven fieldworkers. Graphic 3.4 describes the participation of fieldworkers in dairy recording and their level of experience. Specifically, UK, SL and DN worked in 2004. DE supervised the explorations in 2006 and 2008. Finally, CM, LF and YO assisted DE in the last season. Additionally, there was a number of fieldworkers that excavated at B49 but didn't participate in diary recording, many of them trainees (Eddisford 2008, 356). In general, this evidence indicates a major participation of senior archaeologist in diary recording. However, there were important differences in the recording routine of seniors. Specifically, UK and SL made an entry almost daily, whereas DE, LF and CM made a weekly entry.



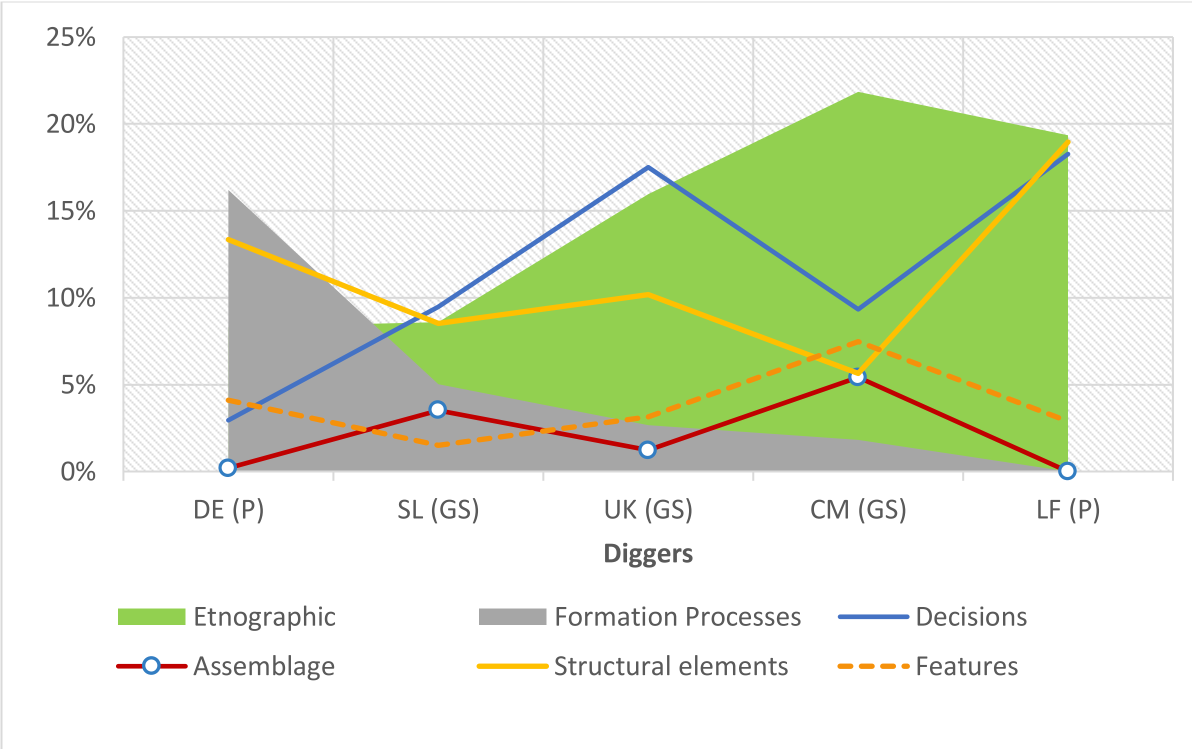
Graphic 3.4. Distribution of the sample by fieldworker, B49. The graphic illustrates the participation of fieldworkers in diary recording and indicates their professional status: professional diggers (P), graduate students (GS) and trainees (t).

The main analysis of the individual performance of fieldworkers was developed by examining the distribution of claims by participant. Graphic 3.5 shows the distribution of a selection of claims that seem to illustrate differences in the performance of fieldworkers produced by skill and background knowledge. For example, the graphic shows an inverse relation in the proportion of formation process and evidential reflexivity with ethnography claims. In other words, the more experienced fieldworkers record formation process and discussions of evidence more frequently, whereas less experienced participant report trivial matters more often. Similarly, there is another important asymmetry in the distribution of spatial and phase data in one hand, and contexts and claims of observational reflexivity on the other. This asymmetry seems to be partially caused by differences in recording routines, hence diggers with higher values of context information and observation reflexivity engaged in a daily routine that produced a more systematic chronicle of discovery. Whereas, fieldworkers with higher values of spatial and phase claims produced a more synthetic record. However, the proportion of spatial and phase data also seems to be higher for fieldworkers that had a longer participation in the excavation of B49 (DE) or participated in the latest phase of exploration (LF), then they benefited more directly from the accumulation of background knowledge. This graphic doesn’t include results for trainees as these were not very significative. DN only had an extensive entry about methodological aspects and YO a short entry describing an episode of layer removal.



Graphic 3.5. Diggers' performance B49a. The graphic illustrates the distribution of various types of claims by fieldworker. These types seem to be significative to indicate differences in the interpretative performance of fieldworkers produced by skill and participation in the project.

Graphic 3.6 illustrates the distribution of claims that seem to be indicative of differences in the type of remains excavated by each participant. This is very clear as one compares feature and assemblage values, for example CM excavated many burials, and therefore less structures (see section 3.5.6). In sum, the analysis of diggers' performance suggests that some areas of recording are determined by skill (formation processes and evidential reflexivity), but others will depend on circumstantial aspect too. For example, a wider background information facilitates the interpretation of spatial and phase data, whereas some aspects are defined by the type of excavated remains.



Graphic 3.6. Diggers' performance B49b. The graphic illustrates the distribtion of different types of cliams by fieldworker. This distribution is considered to be consequence of the type of remains excavated by participants.

Empowerment, authority and responsibility of reflexive diggers

As mentioned before, one of the goals of the reflexive method is the empowerment of fieldworkers for interpretation (Hodder 2005, 652). Yet, the general results of this analysis do not seem to support this idea because the work of B49 diggers is largely confined to the definition of layers and structures (Graphic 3.1). Reflexive archaeology also denounced the low status of fieldworker within a vertical structure od decision-making (Berggren and Hodder 2003, 424). Yet, some entries clearly show that this didn't change very much at Catalhöyük. However, this is a more complex issue because it's also important to remark that experienced staff also can be a source of relevant background knowledge to less experienced fieldworkers as in the following example.

> *Today, we (DGN & SHL) began to dig westwards in an attempt to remove the fill bulk. We stumbled upon several plaster lumps that later Ian and Shanina told us to take out, based on previous experience of these types of house fills. (SL/02/04).*

Moreover, some testimonies of ethnographic, although irrelevant for interpretative issues, they can be a useful about the social relations in the field. Specially, some examples are very revealing of how fieldworkers are still treated as second-rate specialists and how their technical skills can be questioned at any moment. One of the soundest testimonies is related to the discovery of a figurine's assemblage (Example 3.7).

> *Soon, the 'figurines experts' were hovering over me like a protective mother of her younglings. Two of them just sat there for over an hour, cackling, and watched me excavate on the assumption that more figurines would appear. I felt as if I was somehow incompetent. One nameless individual was very upset that I had not photographed the figurines in-situ and that all the meaning of the figurines was lost since I did not know their orientation. He even criticised me that I was excavating improperly and that from this point forward I should continue to excavate with only a brush and a leaf trowel. All rubbish, in my opinion. (SL/25/04).*

Evidently, the former examples only represent particular instances of general problems that have been described elsewhere (Everill 2009; Chadwick 2003). However, an aspect that has not been considered very often are the workload implications of being reflexive. This consequence is very clear in the fragment below that not only indicates that fieldworkers have more duties, which also increment working hours in a dayshift. Then, these aspects should be considered with more detail among reflexive theory. Otherwise, there is a risk of masking labour exploitation as something else. Not to mention, the fragile contractual situation even for some of the experienced fieldworkers being dismissed with apparently no good reason (Farid 2015, 65).

> *IHo's[6] visit at the end of the workday actually provided us with a much better understanding of the interrelationship of the features in space 100. I will not describe everything here, since I still have a lot of unit sheets to enter into the database, but the strategy of the next few days will largely be guided by the insight he gave us. (UK/15/04).*

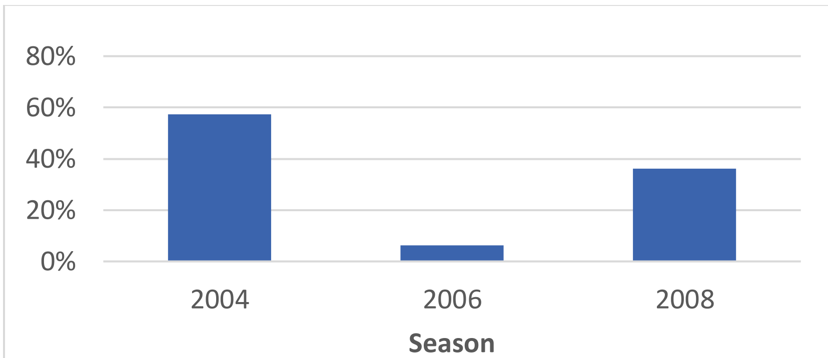
3.8 Longitudinal analysis

The final type of analysis was the longitudinal distribution of claims to investigate similarities and differences between the earliest and latest phases of excavation. This analysis followed the project's seasonal order by year.

---

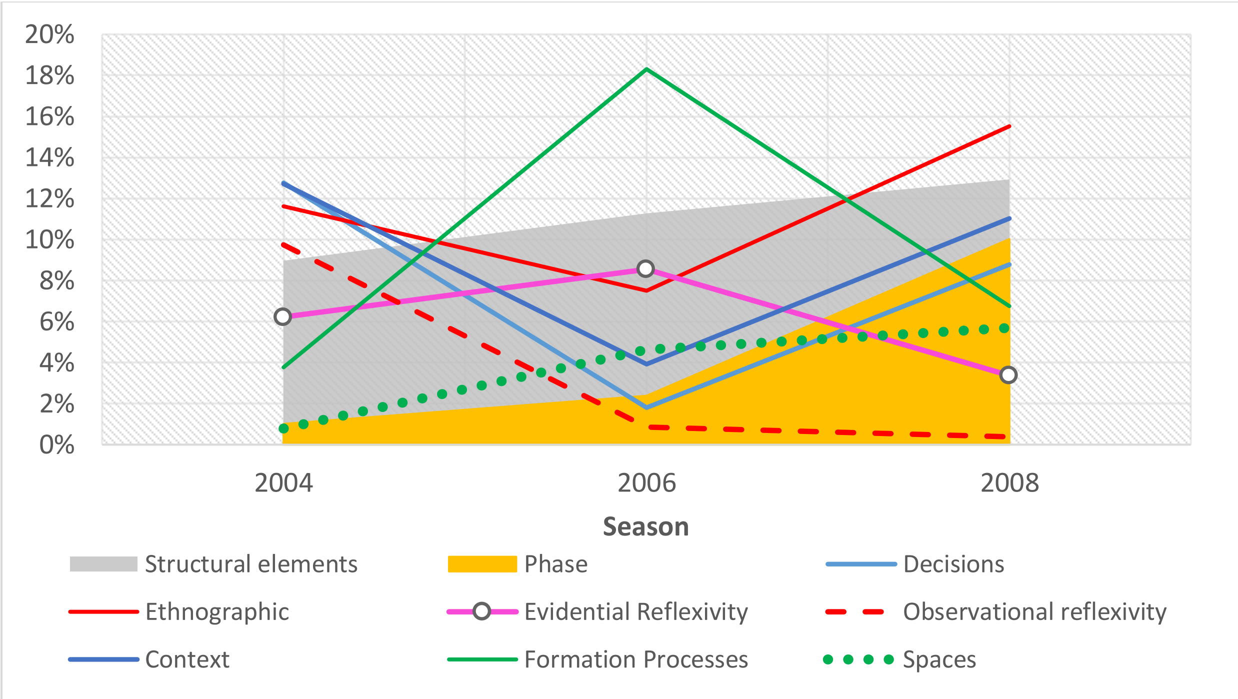
[6] Ian Hodder

Graphic 3.7 shows the distribution of the sample by season, which clearly reflects differences in recording behaviour of diaries. Specifically, the graphic illustrates a more intensive production in the first-year result of a daily routine of diary recording, the second year shows an important decrement in diary production because only one fieldworker recorded a weekly entry. The final year show another increment that represents the production of three fieldworkers making a weekly entry. For this reason, the distribution of claims was calculated independently for each year.



Graphic 3.7. Distribution of the sample by season. B49.

Graphic 3.8 shows the distribution of the most important claims in the three seasons. In general, there are two clear groups of trends. One with higher values in the middle and lower in the extremes (formation processes and evidential reflexivity) and another in which values are higher in the extremes and lower in the middle (ethnographic, decisions and context claims). In general, these results seem to be explained by the recording routine and the skill of fieldworkers. Thus, despite the lower proportion of entries in 2006, this season reports the higher frequency of formation processes and evidential reflexivity. Nevertheless, the most important patterns are the increasing trends for structural, space and phase claims. However, increments in strcutural and spatial data seem to be a direct consequence of the act of discovery that improves the visibility of structures. Besides, this explanation is consistent with the constant decrement of observational reflexivity. In contrasts, the increment in phase claims is better explained by an accumulation of background knowledge in previous years.



Graphic 3.8. Longitudinal analysis. B49. The graphic illustrates the distribution of claims across the seasons of the project.

Finally, the individual and the longitudinal analysis of records provides important evidence of the different scales of interpretive work. In the first place, the analysis by fieldworker shows the diversity of individual performance shaped by skill and participation in the project. In contrast, longitudinal analysis shows more clearly the effects of collective work that creates an accumulation of data that facilitates the interpretation of more complex data over time. However, this collective approach is also very clear in the testimonies of the side room (Example 3.14) and in examples of interactivity during priority tours (Example 3.10 and Example 3.13).

3.9 General observations

The internal analysis of B49 diaries has provided a general view of different aspects such as the recording strategy, the interpretative process and the relevance of reflexive and non-interpretative information. Likewise, the analysis of the performance of fieldworkers and the longitudinal development of the project has provided additional evidence of the incidence of epistemic factors like the skill and the cumulation of background knowledge in the investigation of B49.

The recording strategy. One of the most evident characteristics of dairy entries is their predominant narrative and textual format. However, the most important problem with these testimonies is giving account of many irrelevant aspects. Less frequent are synthetic and argumentative testimonies which constitute relevant pieces of information; however, these data are not systematically recorded and the standard of recording is not always adequate. Less serious is a lack of coordination among fieldworkers to avoid repetitions among entries. Another positive aspect of diaries is providing a good balance of individual and collective practice.

The interpretative process. In general, B49 diaries show an important variability of interpretative data, although most of this information concentrates in the investigation of layers and structures. In general, this is consequence of the narrative strategy and the character of the site. Hence, the excavation of B49 can be described as the exploration of a highly preserved building mainly formed by masonry structures and burials. For this reason, there wasn't much diversity in the investigation of deposits, structures and depositional processes. Moreover, the preservation of remains contributed for the interpretation of complex aspects like constructive phases. Finally, the clearest deficiency in the interpretative process of B49 is failing to address the functional investigation of the building, which was feasible if one considers the epistemic context of the exploration.

Reflexivity. The distinction between observational, persistent and evidential reflexivity was useful to identify relevant reflexive data. This study showed more clearly the importance of argumentative discussions over narratives of discovery. However, the analysis also has shown cases in which relevant reflexive data is not properly reported. Finally, the study of reflexive data showed the positive effects of immediate feedback onsite, which is a powerful mechanism for solving relevant doubts and revising interpretations.

Non-interpretative information. Diaries include relevant information of observational circumstances and operative decisions to contextualize the interpretive practice and the credibility of data. Occasionally, these testimonies are not properly reported, but this is more recurrent when describing sampling decisions. Finally, there is a large component of irrelevant ethnographic information.

Individual and longitudinal analysis. Individual analysis has shown that junior staff has less participation in diary recording. Likewise, the analysis confirmed the incidence of individual skill for some areas of investigation, especially in evidential reflexivity and formation processes. In contrast, longitudinal analysis showed more clearly the incidence of participation and the accumulation of background knowledge for the interpretation of more complex aspects like constructive phases. Altogether, individual and longitudinal analysis showed two faces of field interpretation; individual and collective practice.

Chapter 4 Single context recording

From the moment when we are no longer resigned to purely and simply recording the words of our witnesses, from the moment we decide to force them to speak, even against their own will, cross-examination becomes more necessary than ever. Indeed, it is the prime necessity of well-conducted historical research. Marc Bloch, 1953, 53.

This chapter presents the results of the second case study which examines the registers from a project of commercial archaeology in London recorded with single context recording. Specifically, this case uses as example the explorations at Northgate house identified with the site code MRG95. The chapter begins with a brief account of the history and methodological principles of the recording system, especially its sequential nature (4.1). This is followed by an explanation of the structure of the context sheet and the process of analysis for a sequence of related documents (4.2). Afterwards, there is general description of the epistemic context of the project. In broad terms, the exploration can be characterized as a short-term commercial project executed by a team of professional diggers working in multiperiod site with different levels of preservation (4.3).

The records of MRG95 are stored in the archives of the Museum of London (LAARC) and the chapter also describes some generalities about the sampling process of its sequence (4.4). The longest section describes the qualitative and quantitative results after the internal analysis of MRG95 (4.5). The following section makes some general methodological observations about the importance of documental analysis for evaluating the credibility of testimonies. This is a powerful argument to reject some reflexive ideas regarding the impossibility to assess standardized formats (4.6). The analysis is completed with the analysis of the individual performance of fieldworkers (4.7) and the longitudinal development of the project (4.8). Altogether, these elements will contribute for a clearer understanding of the effects of individual skill and background knowledge the interpretative process of the site. The chapter concludes with a recapitulation of the main observations derived from the analysis of MRG95 (4.9).

4.1 The recording system: a sequence of related events

The invention of single-context recording took place in Britain during the seventies within the context of rescue archaeology. One of the main factors that brought the development of a new recording method was the challenge of recording urban multiperiod sites with deep sequences of stratification (Lucas 2001, 58; Hammer 2000, 144; Spence 1993, 25). In short, SCR was developed for working on large scale projects within stratigraphically complex site and relatively short periods of time. Early rescue projects had been confident in using sections and phase plans, but Harris pointed out later that such procedures slowed the pace of work and raised many problems when recording and interpreting the sequence (Harris 1989, 85–90). Harris and Ottaway reached a solution for this dilemma that consisted in making individual plans of stratigraphic units, that later would be used for interpreting a sequence (Harris 1989, 95). Likewise, when Harris himself worked in the Winchester Research Unit, he devised the Harris Matrix, a diagram to represent the depositional sequence (Harris 1989, 36; Lucas 2001, 56). Hence, this early version of the system commonly referred as single-context planning put more emphasis in its graphic elements than in the context sheet.

According to direct witnesses and second hand testimonies, single-context planning spread over Britain after being adopted by most excavation units at the time (Lucas 2001, 57; Carver 2011, 21). However, the method was largely refined by the former Department of Urban Archaeology in London, especially the written components of the system like a descriptive and interpretative section. This aspect is clearer after comparing the context cards from the seventies and the eighties. In the first case, sheets are mainly designed for single planning and stratigraphic diagrams (see Harris 1989, 36 Fig 38; Spence, 1993, 27 Fig 2.1). It's until the following decade that textual information becomes central. Although, it is in this decade that additional formats like the skeleton sheet appeared too (see Ottaway 1992, 29 Fig 2.4; Spence 1993, Fig 2.2 and Fig 2.3). Equally important is considering further updates in the context sheet, especially for textual data, during the early nineties by the Museum of London (Spence 1993, 33). In this way, the development of SCR can be defined as a twenty-year period that started in the early seventies with a stronger emphasis in graphic elements and culminated in the early nineties with a stronger emphasis in textual information especially layer descriptions and interpretations.

As it is known, the rescue movement pushed the agenda for a full recording of the threated historic environment (Rahtz 1974b). In this context, SCR was a potent tool to accelerate the investigation of urban sites. However, this came with a cost as recording activities were mainly focused in in the definition of layers and interpreting the sequence which remarked the division between field and post-excavation work even more (Hammer 2000, 144). Moreover, many rescue projects required the participation of big crowds given their scale. These people were commonly hired under temporary job schemes to work as excavators (Baker 2011, 196; Everill 2009, 27). However, the introduction of SCR challenged the traditional division of labour between digging and recording staff, incorporating both activities in one person, the excavator (Roskams 2001, 170; Spence 1993, 26). This transitions from labourers to recorders is commonly referred as ‘the empowerment of excavators’ (Eddisford and Morgan 2019, 248; Ottaway 1992, 27; Farid 2000, 24; Hodder 1999, 93). However, this process also brought important social changes as it gave the chance to working classes (usually the target of temporary job schemes) to participate more actively in archaeological investigations, not only as field labourers. In this way, the development and establishment of SCR must be understood as the consequence of multiple factors including the social context created by the rescue movement and the methodological challenges brought by the exploration of multi-period deposits. However, this exciting period of innovation would not have been possible without the economic and political support during the post-war period which contribute for the funding of many excavation units across Britain and contributed to the expansion of the new recording method (Jones 1984; Roskams 2001, 23–29; Ottaway 1992, 11–12; Carver 1987, 104–108).

As explained before, with the introduction of competitive tendering, development led archaeology became suspicious for its economic drive. At the same time SCR became the main target of post-processual critics that particularly questioned its standardized nature, particularly, the rigid format of context sheets. However, this is not a fully accurate view because a context sheet is formed by differ sections and some of them might be less rigid than others. Besides, there is a diversity of formats (deposit, cut, timber, burial and masonry), although in general the structure of all of them is very similar. However, the main problem with critics is reducing SCR to the context sheet and failing to acknowledge that the recording strategy operates as a sequence of related events. Specifically, SCR organizes context sheets into a sequence of natural numbers connected in multiple ways. For instance, by stratigraphic relations and feature relations.

Besides, there is a conditional relation among events because the definition of a layer is based in the definition of previous units. This conditional element shapes the learning process during excavation that contributes to interpretation and reflexivity. However, this is not a lucky coincidence, because this process is regulated by some general principles that should be systematically applied.

The first is a principle of empiricism because it's thought that first-hand experience puts diggers in the best epistemic position to identify and describe layers (Roskams 2001, 170; Spence 1993, 25). Secondly, there is a principle of interpretation. A consequence of direct experience is that fieldworkers are also the best informed about the attributes of layers to interpret its stratigraphic relations, formation processes and so forth. Context sheets include a descriptive and interpretative section for capturing this information. Descriptions provide an account of layer based in its most distinctive attributes, whereas interpretations must explain how observable signs support an interpretation (Roskams 2001, 244). Additionally, there is a principle that can be defined as a reflexive approach to documentation. This principle emphasizes the need to provide "*a descriptive record updated through the excavation process*" (Spence 1993, 25). In consequence, the interpretation of deposits might also change with additional evidence (Roskams 2001, 171). Overall, this principle aims to encourage some sort of reflexive loop between descriptive and interpretative work. Finally, there is a principle of checking up records. When diggers have filled a context sheet, this must be checked by a third party (usually a supervisor) to look for inconsistencies and errors. For instance, if a cut attribute is missing, the supervisor must ask the digger to include this data or explain the absence (Roskams 2001, 235). However, as explained before, SCR presumes that the systematic application of these principles will be achieved when fieldworkers have been trained to perform these activities.

4.2 The analysis of contexts sheets

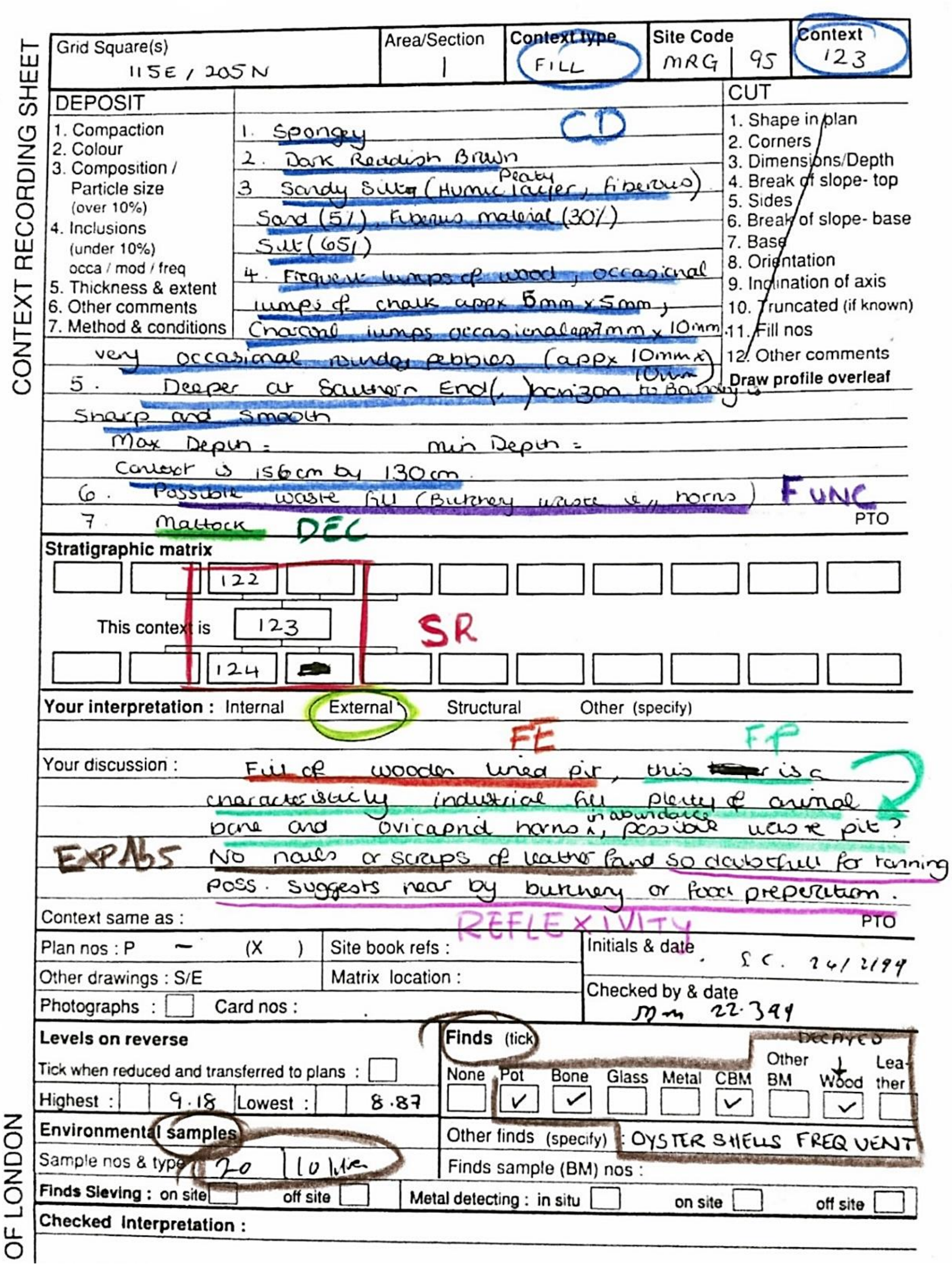
Context sheets are standardized formats organized in different sections to report distinct types of data related to the investigation of layers and cuts. The design and parameters of recording might vary from company to company, but in general all of them follow the same structure. The examined formats correspond to MOLA's system, particularly the updated 1990's version. This includes a collection of formats for describing different types of stratigraphic units: deposits, cut, masonry, timber and burial (Spence 1993; Museum of London 1994). In broad terms, context sheets include a descriptive section in the top, a stratigraphic section in the middle and an interpretive section in the bottom. As explained before, this structure partially guided coding work (see section 2.4). Figure 4.1 presents an example of the most common type of coded document, the deposit/cut format.

More specifically, the context sheet is organized in different sections distributed over the two sides. The upper row of the front side contains referential information (grid, are/section, site code) and identification data (context type and context number). Below is one the main sections, the descriptive box which is includes a list of prompts to describe the main physical attributes of contexts. Additionally, this section includes some prompts to describe observational conditions and decisions (*7. Methods and conditions*), cutting relations (*10. Truncation if known*) and feature relations (*11. Fill nos*), particularly when fills are related to a cut. In the middle there is a stratigraphic section with a Harris Matrix for recording stratigraphic relations and below it's an interpretive section that is organized in two areas: '*Your interpretation*' which contains some predeterminate options to describe spatial aspects (*Internal/External*) and another for *structural* layers. However, the main interpretative area is '*Your discussion*', a free-text section to explain interpretations. In the bottom of this box, there is an additional field for a specific type of stratigraphic relation: correlations (*same as*).

The lowest area of the front side includes more boxes for additional referential data (plan nos, photographs, initials and date of the excavator and check date) and descriptive information: '*levels on reverse*', *find*s and *samples*. Finds indicates the presence and type of recovered finds (*pot, bone, glass, metal, CBM*) or the absence of finds (*none*). Meanwhile, '*environmental samples*' includes a sample number and a short description if a sample was collected. Down in the bottom there is another interpetaive area labelled as '*checked interpetation*', in theory this field is designed to include surpevisor's comments, but the examined record generally included supervisor's corrections over the matrix. Finally, the lowest row includes three boxes (*provisional period, group and initials and date*), these are designed for post-excavation phase. However, some diggers used this area to report cultural/phase claims. The backside of sheets includes two sections; a grid for recording levels on the top and a blank area for drawings in the bottom. This area usually contained diverse types of sketches accompanied of additional textual information occasionally (Figure 4.2). In general, referential and levels information was not coded for being irrelevant for the analysis.

In sum, the format of context sheets facilitated the process of coding. For example, stratigraphy and finds data was coded more automatically, and this was partially true for descriptive areas. However, this task was mor ecomplicated for the interpetive section that often included diverse types of claims (Figure 4.1). Besides, when context sheets included sketches this were coded considering textual information. For example, a context sheets describes a timber structure and therefore the sketch in the back side was coded accordingly (Figure 4.2). However, this also required being aware of the codes of representation for sketches. For instance, a zig-zagged line to represent an uncertain boundary (Figure 4.11). Coding also took account of the relations among context sheet, for example SC-124 (Figure 4.2) describes a timber lining related to a pit-cut HK-134 (Figure 4.3). Then, by comparing the interpretative data in both sheets one can observe the interpretative changes in the investigation of a feature, initially defined as a timber lining but later reinterpreted as a cesspit. The coding process also toook acount of updates and correction on stratigraphic diagramas and sketches made onsite, this was confirmed by studying the handwritting of fieldoworkers. Some sheets included post-excavation annotations identified with a signature and date. In this way, one can separate field and post-excavation information. Post-excavation information was excluded from the analysis.

The following step of the analysis was to assess the credibility of interpretation. In general, this implied checking if documents include an explanation of how evidence supports an interpretation. The best examples usually contained an explanation in the interpretative section of sheets (Figure 4.1). However, some cases required considering different bits of dispersed information to follow the rationale of interpretation (Example 4.12). Likewise, checking the validity of some interpretations required examining the content of two or more sheets. For example, cut sheets usually summarized the interpretation of features considering the evidence from various layers (see section Features4.5.8). This analysis allowed to classify interpretive claims in two groups: valid interpretations and invalid interpretations to differentiate between justified and unjustified claims respectively. In sum, even if the format of the context sheet facilitates some aspects of the analysis, this always required a contextual approach.



CONTEXT RECORDING SHEET

OF LONDON

Grid Square(s): 115E / 205N | Area/Section: 1 | Context type: FILL | Site Code: MRG 95 | Context: 123

DEPOSIT
1. Compaction
2. Colour
3. Composition / Particle size (over 10%)
4. Inclusions (under 10%) occa / mod / freq
5. Thickness & extent
6. Other comments
7. Method & conditions

CD

1. Spongy
2. Dark Reddish Brown
3 Sandy Silty (Humic layer, Peaty fibrous)
Sand (5%), Fibrous material (30%)
Silt (65%)
4. Frequent lumps of wood, occasional lumps of chalk appx 5mm x 5mm, Charcoal lumps occasional appx 7mm x 10mm
very occasional rounded pebbles (appx 10mm x 10mm)
5. Deeper at Southern End (, horizon to Boundary is Sharp and Smooth
Max Depth: min Depth:
Context is 156cm by 130cm.
6. Possible waste fill (Butchery waste i.e. horns) FUNC
7. Mattock DEC

CUT
1. Shape in plan
2. Corners
3. Dimensions/Depth
4. Break of slope- top
5. Sides
6. Break of slope- base
7. Base
8. Orientation
9. Inclination of axis
10. Truncated (if known)
11. Fill nos
12. Other comments
Draw profile overleaf

PTO

Stratigraphic matrix

122

This context is 123

124

SR

Your interpretation : Internal External Structural Other (specify)

FE FP

Your discussion : Fill of wooden lined pit, this is a characteristically industrial fill plenty of animal bone and ovicaprid horns in abundance, possible waste pit?

EXPAbs No nails or scraps of leather and so doubtful for tanning, poss. suggests near by butchery or food preparation.

Context same as :

REFLEXIVITY

PTO

Plan nos : P – (X) | Site book refs : | Initials & date SC. 24/2/94
Other drawings : S/E | Matrix location : | Checked by & date MM 22.3.94
Photographs : | Card nos : |

Levels on reverse
Tick when reduced and transferred to plans :
Highest : 9.18 Lowest : 8.87

Finds (tick) DECAYED
None | Pot ✓ | Bone ✓ | Glass | Metal | CBM ✓ | Other BM | Wood ✓ | Leather

Environmental samples
Sample nos & type 20 10 litre

Other finds (specify) : OYSTER SHELLS FREQUENT
Finds sample (BM) nos :

Finds Sieving : on site off site | Metal detecting : in situ on site off site

Checked Interpretation :

Figure 4.1. Frontside of a context sheet describing the interpretive process of a fill of a wooden lined pit (SC-123). The image also illustrates the coding process of a context sheet into different types of claims.

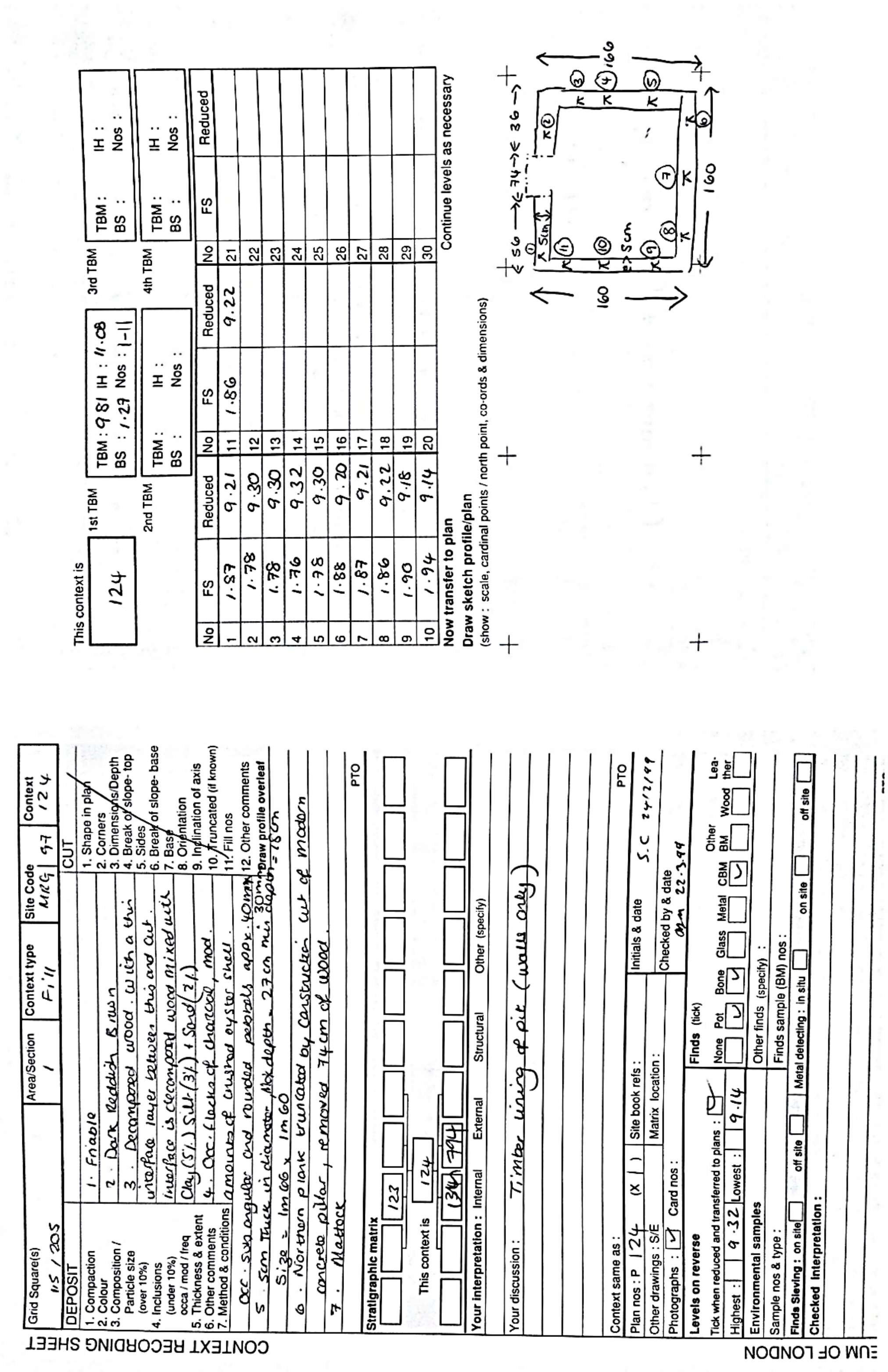
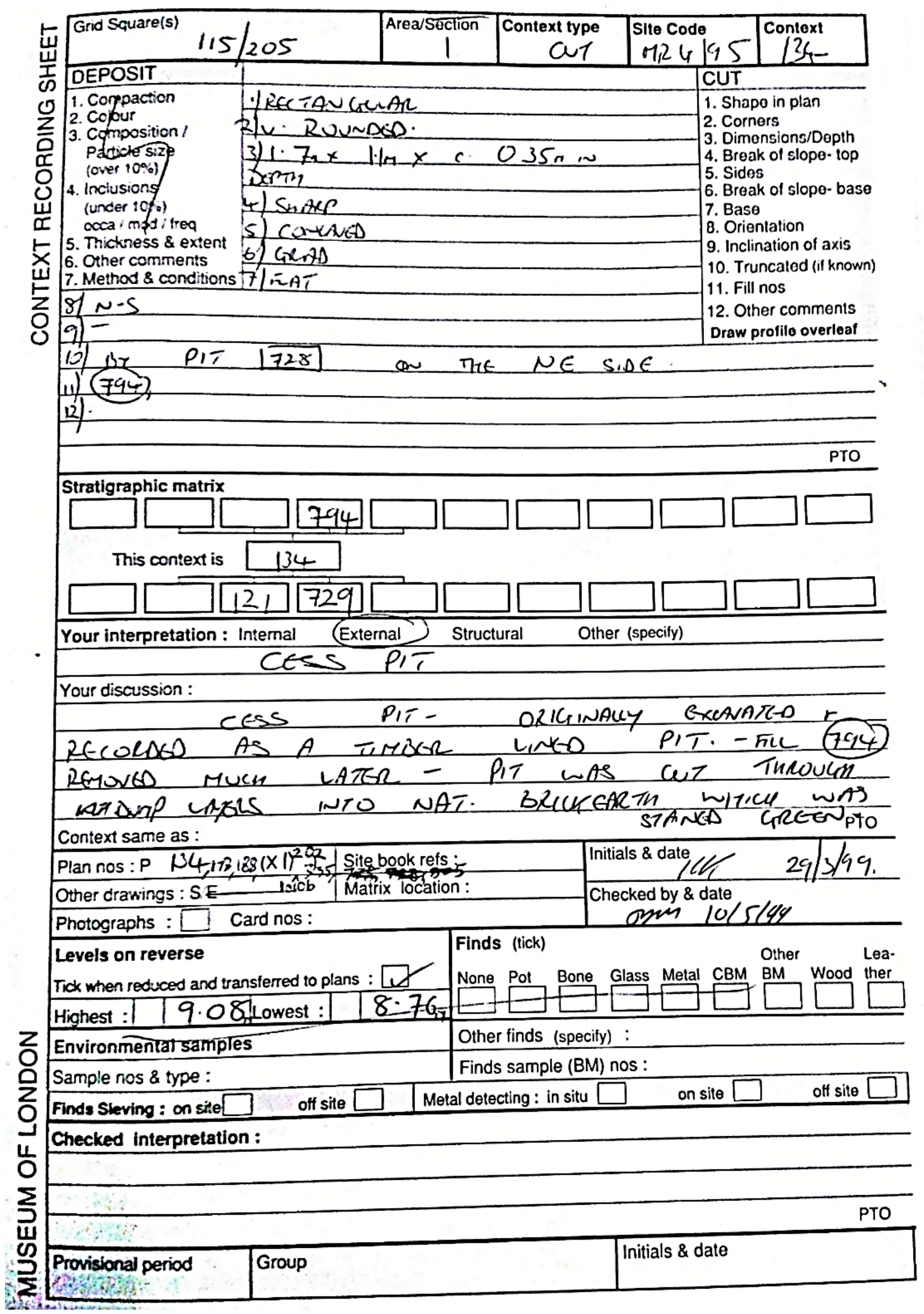


Figure 4.2. The front and back side of a context sheet describing a timber lining (SC-124).


CONTEXT RECORDING SHEET

MUSEUM OF LONDON

Grid Square(s) 115/205 | Area/Section 1 | Context type CUT | Site Code MRG 95 | Context 134

DEPOSIT
1. Compaction
2. Colour
3. Composition / Particle size (over 10%)
4. Inclusions (under 10%) occa / mod / freq
5. Thickness & extent
6. Other comments
7. Method & conditions

CUT
1. Shape in plan
2. Corners
3. Dimensions/Depth
4. Break of slope- top
5. Sides
6. Break of slope- base
7. Base
8. Orientation
9. Inclination of axis
10. Truncated (if known)
11. Fill nos
12. Other comments
Draw profile overleaf

1) RECTANGULAR
2) V. ROUNDED.
3) 1·7m x 1m x c. 0.35m in DEPTH
4) SHARP
5) CONCAVED
6) GRAD
7) FLAT
8) N-S
9) —
10) BY PIT 728 ON THE NE SIDE
11) 794
12)

PTO

Stratigraphic matrix

794

This context is 134

121 | 729

Your interpretation : Internal External Structural Other (specify)

CESS PIT

Your discussion :

CESS PIT - ORIGINALLY EXCAVATED & RECORDED AS A TIMBER LINED PIT. - FILL 794 REMOVED MUCH LATER - PIT WAS CUT THROUGH MADEUP LAYERS INTO NAT. BRICKEARTH WHICH WAS STAINED GREEN PTO

Context same as :

Plan nos : P 134, 172, 183 (X 1) 292/355 | Site book refs : | Initials & date 29/3/99.

Other drawings : S E 136b | Matrix location : | Checked by & date 10/5/99

Photographs : | Card nos :

Levels on reverse

Tick when reduced and transferred to plans : ✓

Highest : 9.08 Lowest : 8.765

Finds (tick)

None | Pot | Bone | Glass | Metal | CBM | Other BM | Wood | Leather

Other finds (specify) :

Finds sample (BM) nos :

Environmental samples

Sample nos & type :

Finds Sieving : on site off site | Metal detecting : in situ on site off site

Checked interpretation :

PTO

Provisional period | Group | Initials & date

Figure 4.3. The front side of a context sheet describing a cesspit cut (HK-134).

4.3 The project and the site: MRG95

The second case study covers a short-term commercial project in Northgate house, London and identified with the site code: MRG95. The site was located at 20-29 Moorgate Street (Figure 4.4) and archaeological services were provided by the Museum of London, (then known as MoLAS). The excavation methodology was based in single-context recording following the protocols defined by MOLA (Museum of London 1994). The project was structured following the principles of MAP2 with an excavation and post-excavation phases (English Heritage 1991). The excavation lasted for 15 weeks between the 12 of February and the 28 of May of 1999. Fieldwork was performed by professional diggers also known as site-assistants and supervised by two senior archaeologists (Seeley and Drummond-Murray 2005, 1). The project included a desk-based assessment and evaluation that determined the industrial character of the site. It was located in the Walbrok Valley, an area formerly known as the industrial sector of Roman London (Seeley and Drummond-Murray 2005, 5). Hence, the excavation had the purpose of investigating the range of industrial activities at MRG95 (Museum of London 2000, 15–16). The excavation results and post-excavation assessment determined the relevance of the site for publication and a few years later, a monograph was published (Seeley and Drummond-Murray 2005).

Figure 4.4. Site location in relation to Roman London map (Copyright MOLA).

MRG95 was a multi-period site ranging from prehistory to post-medieval periods. Yet, the main occupation dated from the roman period. This included a pottery complex, defined by some kilns (S2, S3) and storage pits in located in the centre of the site (Figure 4.6). The site also included an access road (R1) and some structures and features related to the course of the river (S10) (Figure 4.5). In general, the archaeological deposit included different types of layers and features with different level of preservation. The pottery kilns (S2 and S3) were among the best-preserved elements (Figure 4.19) whereas further evidence was more fragmentary.

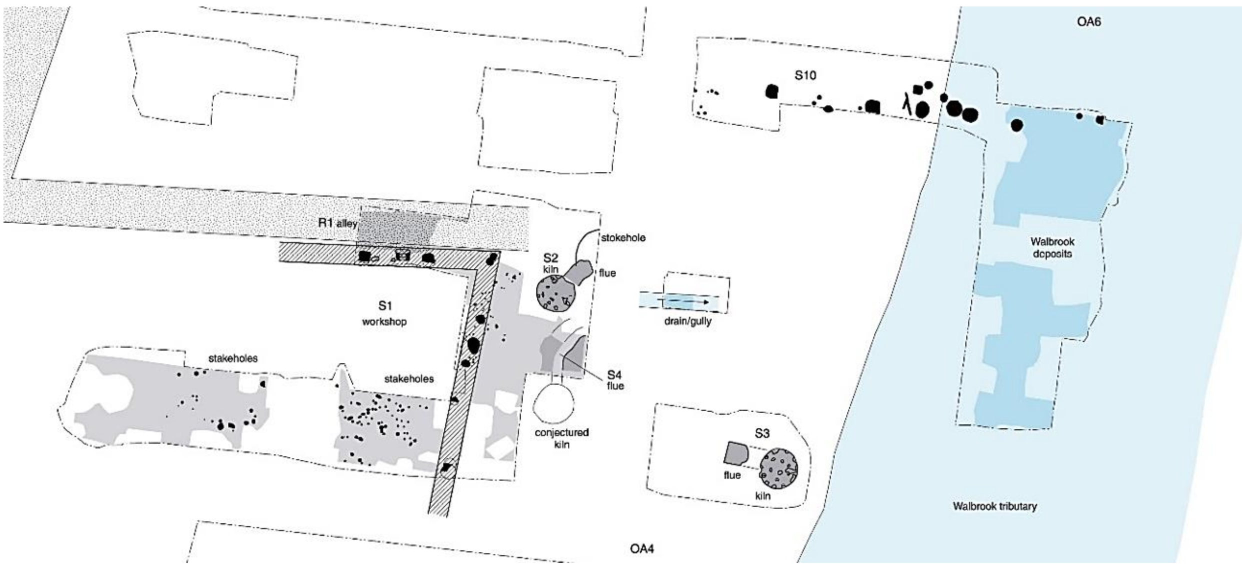


Figure 4.5. MRG95. Furnace complex. Roman period, Phase 1 (Copyright MOLA).

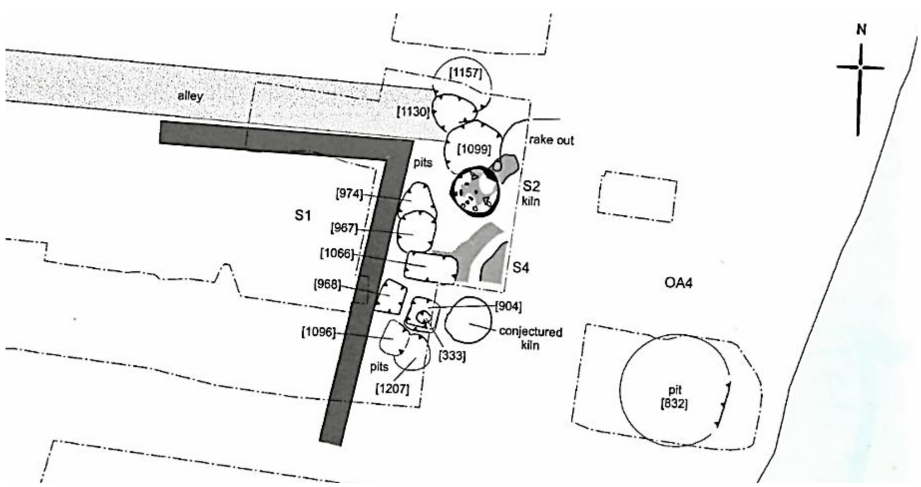
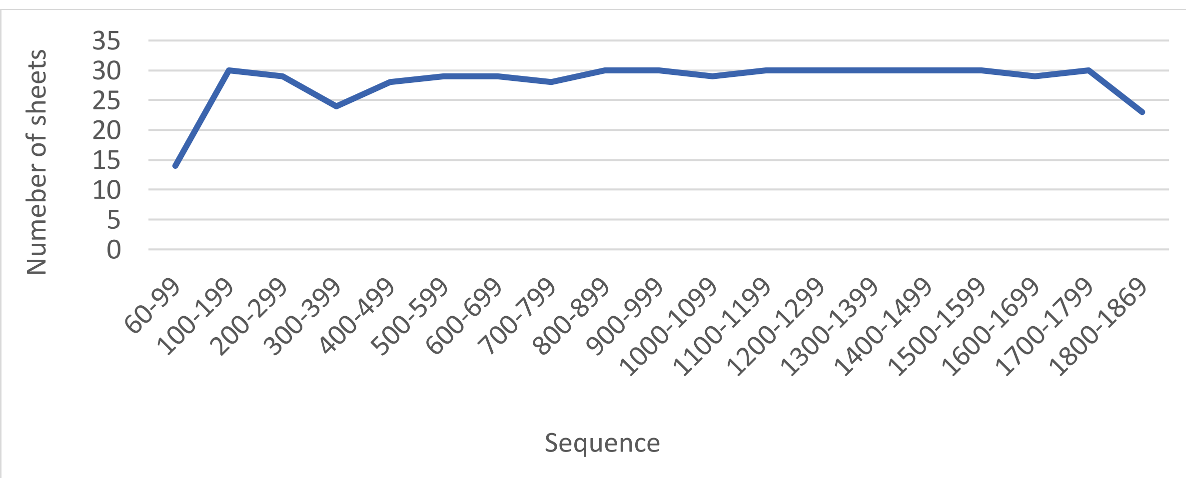


Figure 4.6. Nuclear area of furnace complex. Roman period, Phase 1 (Copyright MOLA).

4.4 The sample

The archive of MRG95 is stored at the London Archaeological Archive (LAARC). The excavation sequence includes 1776 sheets starting in context number 62 to 1864. The previous 61 sheets correspond to evaluation work and were not considered in this study. The sequence was organized in sub-sequences of hundred according to natural numbers {100…199, 200-299}, except for the first and last segments being shorter {61…99} and {1800…1864}. A sample of 30% of the excavation sequence was analysed, this included 532 context sheets. The sample was defined by a strategy that selected approximately thirty sheets by each hundred. Sheets always were collected in groups of five to ten continuous sheets {82, 83, 84, 85, 86}. This strategy was designed to collect related sheets, for example context from a feature or stratigraphically related layers. A random selection process was avoided because the sequence wasn't considered a series of independent events. The sample also aimed to be representative of different type of remains and areas of the site, therefore a previous consultation of the archive and a context list provided by MOLA helped for this purpose. Altogether, these strategies contributed for having a representative sample of various participants and phases in the project. Initial sampling work took place at MOLA's headquarters in July 2017, digital copies of context sheets were taken with a standard mobile phone and processed with an application to improve image quality.

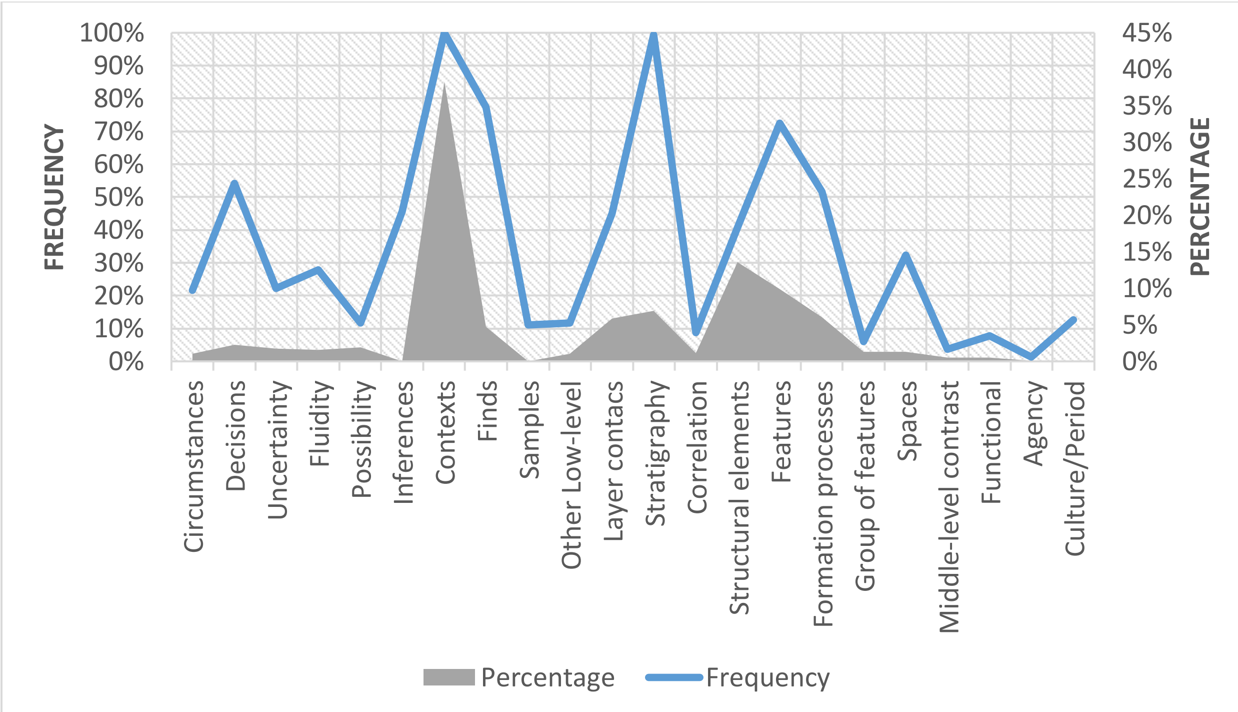


Graphic 4.1. Distribution of the sample. MRG95. The graphic illustrates the number of sheets collected by each hundred of the site's sequence.

4.5 Results

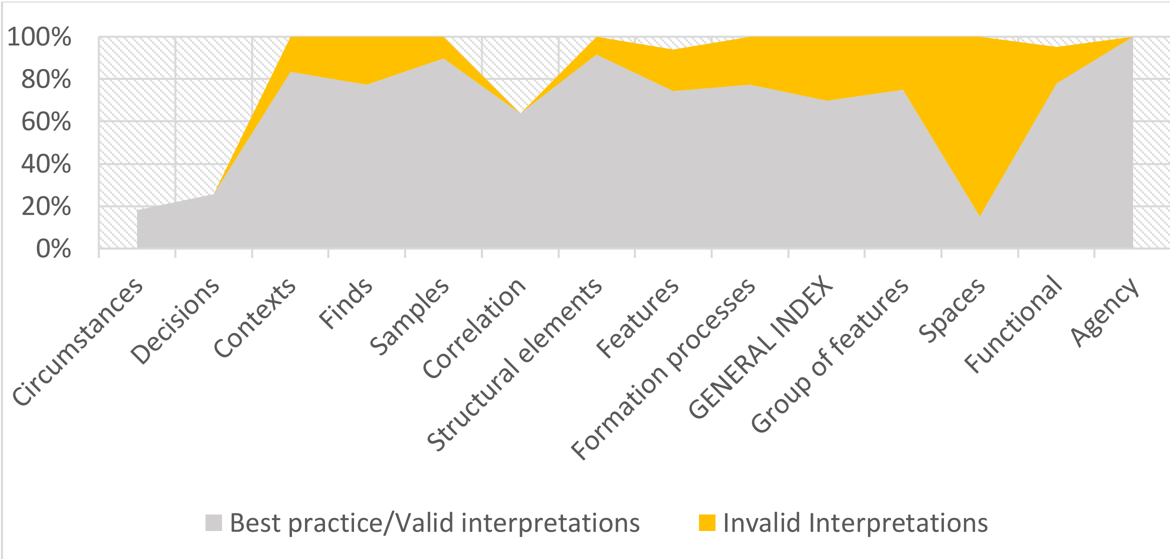
The documental analysis of MRG95 implied the analysis of a longer sequence containing textual and visual data. These characteristics had some effects in the quantitative data of MRG95 which includes frequency and percentages, unlike B49's analysis that only provides percentages. Frequency indicates the recurrence of a type of claim within the sample (how often), whereas percentage indicates the proportion of each type in relation to the total coded information (how much). Frequency was preferred for the internal analysis of MRG95, but percentage will be used for the comparative analysis. Graphic 4.2 illustrates the distribution of claims in MRG95 which shows a similar trend for frequency and percentage values. Overall, this graphic shows a large variability of information, including non-interpretive and reflexive claims. However, context and stratigraphic information are among the most frequent, these are followed by structure, feature and formation process claims. And everything else is less common. Assemblages were not very frequent in the site and only was burial was sampled, this explains the low frequency of this data, therefore no further analysis of this aspect will be given. Building and phase claims were absent from site records too, however these absences deserve further analysis.

A recurrent a characteristic of context sheets was the abundance of retrospective inferences and Graphic 4.2 includes the frequency of these conjectures. Technically, these are not a type of claim because an inference can be functional or depositional. However, the frequency of inferences was an indicative evidence when examining the performance of fieldworkers and the longitudinal development of the project. This is the main reason for having frequency of inferences but not percentage. Similar asymmetries are observed for finds, stratigraphic and spatial data. However, these are explained by the format of recording. For example, stratigraphy is present in almost every sheet in a matrix Harris, but the percentage of coded information is not very high.



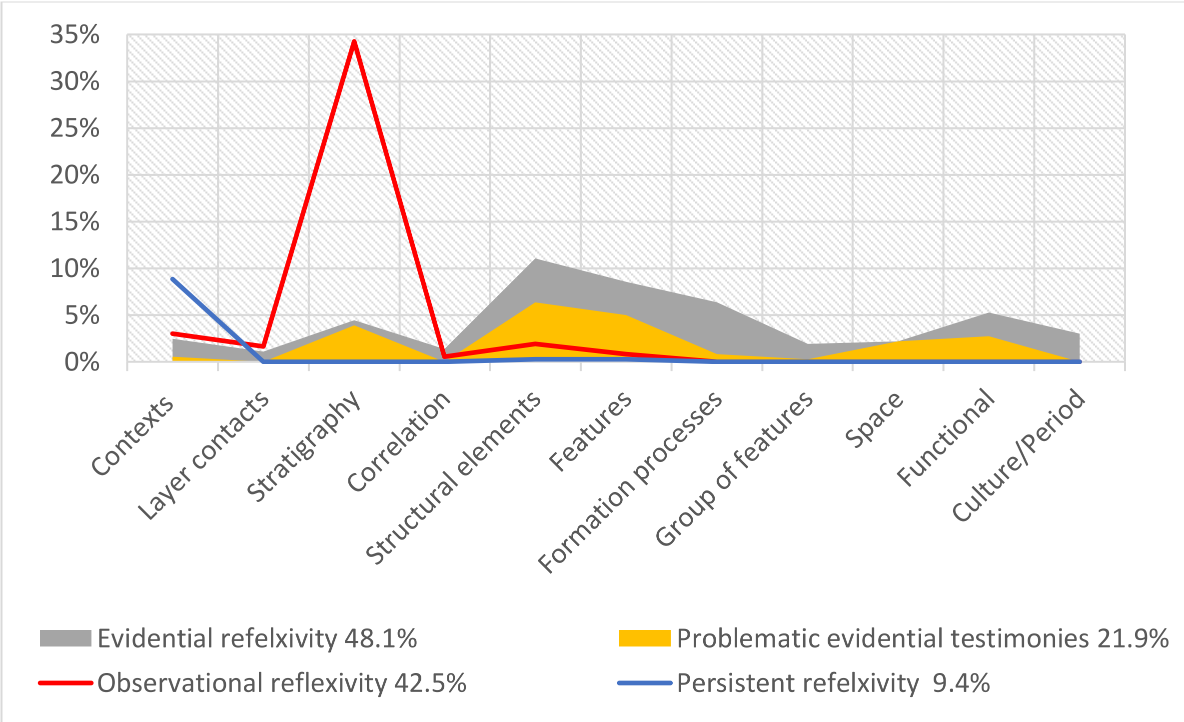
Graphic 4.2. General distribution by type. MRG95. The graphic illustrates the frequency and percentage by type of claim. Frequency indicates how often a claim was recorded and percentage indicates how much information of a type was coded.

Graphic 4.3 shows the results for the credibility of interpretive information. Valid interpretations indicate the proportion of reliable information or justified data. Invalid claims represent cases of unreliable data. The percentage of valid and invalid claims was calculated independently for each type. There are some columns when the sum of valid and invalid claims doesn’t reach 100% (Correlation, Feature and Functional) because some claims couldn’t be evaluated. This graphic includes an additional column labelled as General Index, it represents a general assessment of the credibility of the sample based in the proportion of valid and invalid context, structural and formation processes data. The logic of these shall be explained along the following sections. The Graphic also indicates the proportion of Best Practice Circumstantial and Decision claims, this highlight the most relevant non-interpretative data without implying that complementary information was unreliable.



Graphic 4.3. Credibility of interpretative claims. MRG95. This graphic describes the distribution of valid claims and invalid claims for interpretaive information. It also includes the proportion of best practice tetsimonies for non-interpretative data (circumpstances and decisions).

Finally, Graphic 4.4 illustrates the distribution of reflexive claims in relation to interpretative types and according to the classification of reflexive claims into trivial, persistent and relevant data. Overall, this chart shows an even distribution between trivial and relevant claims. However, as it will be explained, most cases of trivial reflexivity are coded from corrections in stratigraphic matrixes and drawings. Meanwhile evidential reflexivity was coded from textual information. Additionally, there is a small amount of persistent reflexivity related to layer descriptions. Finally, this chart also indicates the number of problematic testimonies for evidential reflexivity.



Graphic 4.4. Reflexivity. MRG95. The graphic illustrates the distribution of reflexive claims into three categories (observational, persistent and evidential) and its relation to interpetive aspects. Also, it illustrates the frequency and distribution of recording problems for evidential claims.

4.5.1 Non-interpretative information

Two types of non-interpretive information were found. Information about observational circumstances was distributed over 20% of the sample (Graphic 4.2). Commonly, this information was recorded in the section for 'methods and conditions' indicating that excavation was done under artificial light and indoors (Figure 4.1). Although this was not systematically recorded, such information provided a general picture of the observational conditions onsite, namely "*a regular unchanging basement environment*" (RC-82)[7].

[7] An identification code is used here and onwards for referencing testimonies of MRG9's context sheets. The code includes the initials of the author and the context number.

Figure 4.7. Observational conditions at MRG95 (Copyright MOLA).

One fifth of these sheets included additional comments and/or sketches to describe specific situations in more detail. For instance, when a layer/feature was partially observed due the limits of excavation (Figure 4.8). In general, this partial observation didn't have important interpretative consequences, except for the extent of layers, which sometimes was emphasized in context sheets: "*Extends beyond the limits of excavation to the north*" (HK-322). Another recurrent example was when a layer or a feature was exclusively observed in section, and therefore context sheets described this circumstance: "*Machine left large section showing the posthole was narrow…*" (FC-64). Altogether, these short descriptions were classified as instances of best practice, being distinguished from the most systematic information (Graphic 4.3).

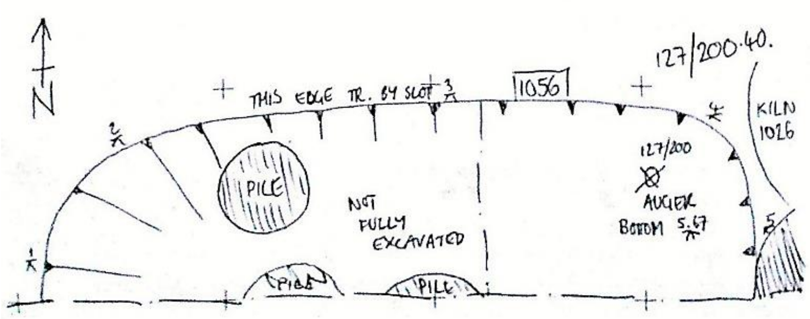


Figure 4.8. Sketch from a partially excavated cut (MW-1056).

Decision claims covered above 50% of the sample (Graphic 4.2). Once more, these sheets included systematic information indicating whether a layer had been excavated with trowel or mattock (Figure 4.1). Overall, this information was not very important, except as a general account of technical procedures. However, one fifth of these sheets included additional comments explaining local operative decisions. These decisions can be organized in two groups, those with interpretative impact and those with no impact. The first group includes examples of technical actions as when someone "*accidentally overcut*[s] [a layer] *during excavation*" (HLB-1047) or when "*mattocking of area may have obscured the full extent of stakeholes [753]; they definitively truncate [546]*" (JA-542). In general, these claims help to contextualize the reliability of data (Figure 4.9). Still, it's important to understand that not every technical accident produces interpretive problems. For instance, one sheets annotates when a wood plank from a well was "*broken in excavation*" (JA-1806) but the interpretation of the layer was never compromised by such action. Likewise, technical actions not always describe mistakes and some sheets describe specific procedures for special purposes like: "*excavated at the end by machine to determine depth of natural in this area*" (FAP-1343). Decision claims also included descriptions of recording actions based in interpretative decisions, for example recording various layers as one context or slicing a layer into various arbitrary levels. Likewise, sheets also included sampling information, but these will be examined later.

Operative and recording decision with no interpretive impact include testimonies describing pragmatic choices due multiple factors. For instance, the partial exploration of a structure due safety protocols as explained in one sketch: "*true extend of cut difficult to ascertain in east due to danger of concrete collapse, therefore extend of well sough instead*" (JA-1759) (Figure 4.9). This group also includes recording decisions without affecting interpretative aspects: *"essentially this context was numbered as [166] and planned as such. This number kept for finds*" (MLC-187). Overall, these additional comments are important to clarify technical procedures without affecting the credibility of interpretative data.

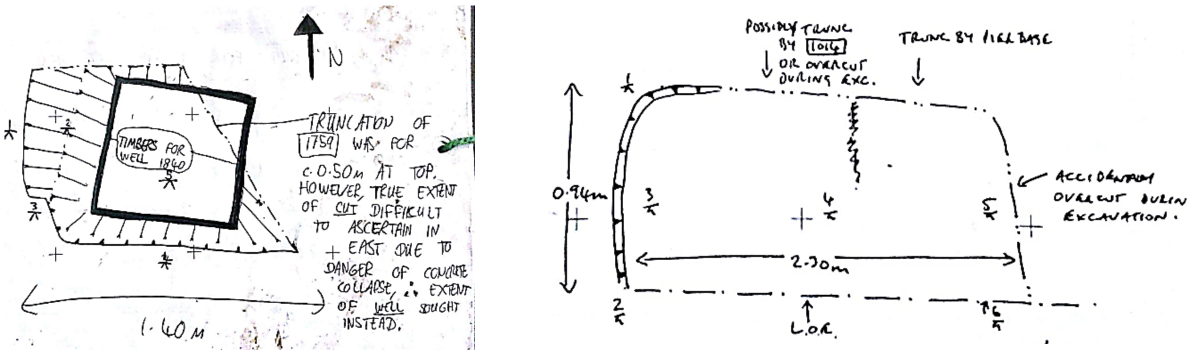


Figure 4.9. Sketches describing operative decisions. JA-1759 (left) and HLB-1047 (right).

Altogether, such testimonies of operative decisions were counted as instances of best practice, unlike the most systematic information that only describes if a layer had been trowelled or mattocked (Graphic 4.3). This doesn’t imply that systematic information is incomplete or redundant. More exactly, best practice circumstantial and decision data exemplify cases when context sheet include additional information to contextualize the observational or the recording process of something, and more importantly to contextualize the credibility of interpretative data when required. Thereby, context sheets use supplementary comments and sketches to describe relevant background information which fulfils the same function that relevant non-interpretative content in reflexive diaries.

4.5.2 Contexts

Context descriptions are the most frequent claim in the sequence (Graphic 4.2). This includes layers and cut descriptions based in the systematic account of physical attributes. Besides, some of these sheets include complementary sketches depicting the shape of units. Furthermore, a closer look at layer descriptions revealed two descriptive patters. In general, context descriptions provide an abstract account of layers and cuts, based in a systematic description of attributes. However, descriptions of attributes are frequently subject to some form of idealization; for example, chromatic variability is summarized into a general statement like ‘dark reddish brown’ (Figure 4.1). Additionally, there are some sheets when fieldworkers adopt an alternative approach and they provide a more exhaustive description of some attributes. For instance, one sheet observes compositional variations in a burnt clay layer: “*context here more mixed, does contain burnt brick clay fragments but also greyish clay*”, unlike the other half that included "*burnt brick clay debris more concentrated east*" (FC-163). However, the most interesting examples is when such exhaustive description became crucial for the interpretation of layers: "*Much slag in western half and pot all over…suggests this was an attempt to consolidate the area specially the western half*".

Moreover, the analysis of layer descriptions provided some examples of alternative mechanisms of layer definition. One of the commonest procedures was the use of arbitrary levels and generally, testimonies are very clear to explain the rationale of such operative choices as in the following example: "*number attributed to top portion of fill, which was separated at an arbitrary point to ascertain whether any difference in age of pottery was discernible*" (LMC-511). Less often, diggers recorded multiple layers as one context like in the example below. This testimony provides some clues about the rationale of decision (the thinness and similarity of layers) but is less successful to explain additional factors such as the technical difficulty to remove thin layers which probably explains the type of sample.

Example 4.1

> *"Fill of ditch [546]. Fill was made up of many deposits, with finely sorted alluvial deposits towards the bottom, at the base of [546]. A column sample was taken of the alluvial layers […] excavated as one fill." (HLB-540).*

Another similar example was the definition of context [1424] from a group of horizontally dispersed deposits. This decision was explained with a sketch and instead of describing each deposit as an individual unit, the contexts includes all those bits due its similar nature described as 'burnt area' (Figure 4.10). Overall, these examples portray a variability of interpretive and descriptive procedures that are carefully explained in context sheets. Likewise, such variability indicated that the rules of layer definition and description can be adjusted to the interpretive needs of fieldworkers. In other words, the system is not fully rigid as critics assert.

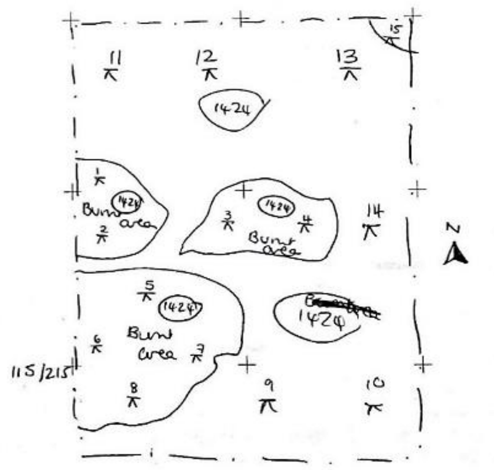


Figure 4.10. Sketch of a layer (SC-1424).

In addition to layers description, context sheets must indicate the type of layer being described, namely whether it's natural, a dump, a masonry deposit and so on. Sometimes this information was included in the section for context type, but more often it was reported in the interpretive section (Figure 4.1). Nevertheless, one must be aware that one thing is describing a type of context (layer/deposit, fill, cut); and another defining the type of layer (natural, structural, midden, dump). Based in this distinction a group of context sheets was identified which included a type of context (layer/fill) but failed to stablish a type of layer. Many of these sheets only included descriptive information within the interpretative section like in the following examples: "*burnt clay layer*" (L'ON-202) or a "*silty clay rich layer*" (GT-883). Therefore, the main problem with these testimonies is a lack of interpretation or failing to define the type of layer being described (see Figure 6.1). In consequence, these sheets were classified as instances of invalid context data. In general, this type of problem was more recurrent for sedimentary layers. Cuts sheets normally mention the type of cut being described (post-hole, ditch and pit, for example) which generally are considered cultural events. Only when a cut was considered natural, its type was remarked: "*probably tree/root disturbance*" (LMC-224). In sum, around 80% of the sheets included a reliable description of the type of layer/cut being investigated and therefore these examples were classified as instances of best practice (Graphic 4.3). .

Context sheets include uncertainty and fluidity information when defining layers. The most recurrent example are interpretive doubts in layer boundaries. Generally, these cases were graphically represented with a zig-zagged line representing an ambiguous area (Figure 4.11)

Occasionally, this could be complemented with an explanatory comment describing the ambiguities of evidence: "*horizontal extend indeterminate as merges with clay earth and pebble layer to the north*" (MLC-856). But in none of these cases, uncertainty compromised the confidence of layer definition, this was clearly explained in one sheet that mentions: "*feature easily distinguishable, but outer region to east less certain as distinctly lighter (dotted area on plan)*" (MLC-1030) (Figure 4.11). Hence, these examples were classified as instances of persistent reflexivity in context definition (Graphic 4.4).

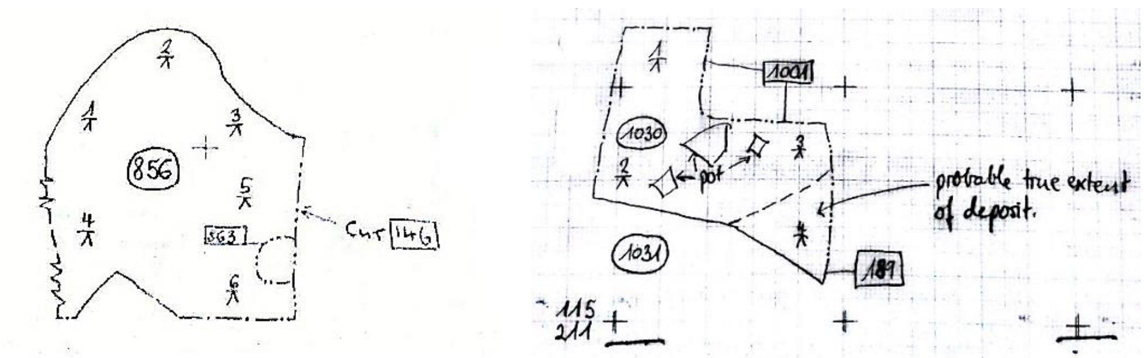


Figure 4.11. Sketches depicting ambiguous layer boundaries. MLC-856 (left) and MLC-1030 (right).

There were only a few cases when evidence produced some doubt in the existence of layers as in the example below. Similarly, in another case a sheet describes a retrospective reconsideration regarding the number of layers being defined: "*On excavation this layer found to be smaller than thought. A very gravelly and cobbly area of it given a separate number [1262] Similar to 1262 but less stony*" (FAP-1109). These cases were defined as evidential reflexivity and most of these examples included short discussions of evidence, except for a few cases that were considered instances of problematic reflexive recording (Graphic 3.3).

> *"Possible fill of a cut feature/However no cut edge was obvious […] The surrounding fill of a borehole intruded into the area and obscured half of the area where the possible interface between the two would have showed up" (FAP/JAP-1034).*

Finally, the sample included a few examples of observational reflexivity in context definition. Some sheets include brief notes describing uncertainty at early stages of discovery. For instance, a sheet mentions: "*Not clear whether layer or fill*" (JDM-195), but later it became evident that such layer was part of a kiln structure. In another case, when excavating a big fill, a context sheet mentions: "*possibly extending over the other side of foundation trench, yet to be excavated*" (FAP/JA-517). Then, when the excavation continued, the additional portion was recorded with a new context number. In general, these are uncommon testimonies that describe temporary doubts when excavation was paused for a reason.

4.5.3 Finds and samples

Finds information was present in 80% of the sample (Graphic 4.2). This data was generally located in the finds section indicating the type of recovered finds or absence of finds, but often this data was also included alongside layer descriptions within inclusions (Figure 4.1). The remaining 20% of sheets didn't include this information and therefore represent a minor recording failure (Graphic 4.3). When required, sheets also mention relevant aspects of the recovery process: "*pottery recovered mostly from the horizon with layer below*" (LAB-1546), or when: "*a barrel shaped part of well… left in situ. To dangerous* [to collect]" (JA-858). Unlike the previous case study, no instances of reflexivity related to finds detection was observed for MRG95.

Around 10% of the sheets reported sampling activity (Graphic 4.2). Normally, this was indicated with a sample number and a type of sample (Figure 4.1). In general, sheets didn't include explanatory comments for these decisions. Nevertheless, interpretive information usually provided relevant context for understanding sampling actions which can be classified as follows. Some of layers were sampled due their rich assemblages, this included fills that "*contains a lot of leather*" (FC-76), layers with rich assemblages of pottery, slag or organic materials then requiring an "*environmental sample taken as high-level of burnt material present and some badly decayed wood*" (JA-1622). Likewise, fills inside fire-features (ovens and kilns), surface layers, post and stake fragments were systematically sampled due its interpretative potential (Figure 4.12). Most of these sampled were taken following the policy defined by the company (Museum of London 1994).

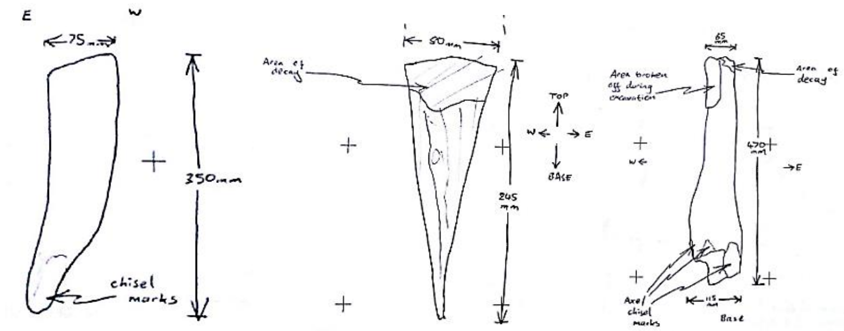


Figure 4.12. Sketches of timber stakes. GT-1227 (left), GT-1249 (centre), GT-1259 (right).

Exceptionally, some sheets included additional information to explain sampling decisions, for example when a layer hadn't been properly explored onsite: "*deposit remained unexcavated, but a sample was taken*" (HLB-1182) (see also Example 4.1). Likewise, sheets included explanatory comments when an interpretative claim was unsure, and sampling ensured further study: "*Possibly burnt clay lining. Occurred mainly to NW of pit near the edge. Almost lining the pit. All pieces sampled, retained*" (GT-1411). Overall, these clear examples where one can appreciate the importance of documental analysis and despite the lack of a diary testimony, the logic of sampling decisions usually can be understood by cross-examination of sampling and interpretive information. Only 10% of the sheets could not be understood due the lack of interpretive information, therefore being counted as unreliable sampling information (Graphic 4.3).

4.5.4 Other Low-level claims

Other low-level claims summarize the frequency of low-level contrasts, expected absences and indicative absences. The frequency of these claims is very low (Graphic 4.2), however they represent important data within site records. Low levels contrasts include a few examples when comparing the attributes of two or more layers or finds. Yet, the most interesting examples are when such observations are used to solve interpretive dilemmas: "*A very gravelly and cobbly area of it given a separate number [1262] Similar to 1262 but less stony*" (FAP-1109). Likewise, low-level contrast are important observations to justify correlations (see section 4.5.6). Cases of expected absences describe situations when an expected item was no observed, for example when describing a fill from a kiln, a record mentions: "*does not contain a lot of pot*" (MW-931). Yet, the most frequent and important claims can be useful to provide a reflexive testimony. For example, to evaluate the strength of an explanation in the form of a counterfactual argument. Then, a sheet describing a keyhole cut mentions: "*cut possibly industrial, but sides not burnt, if it's a small hearth*" (TM-127). Finally, the sample included a few sheets with instances of indicative absences or when the absence of something is considered evidence of something else: "*flue appears to be tunnelled as there is no clear cut visible after removing furnace lining*" (MW-545). This also included a few cases of negative evidence like an imprint in the fill of stake holes: "*negative traces of organic wood*" (SC-608). A common characteristic of these claims is that they can only emerge in the light of previous experience which enables to compare things and have expectations.

4.5.5 Stratigraphic information

Stratigraphic information includes three types of claims, layer contacts, stratigraphy and correlations. Stratigraphy is the most frequent piece of information systematically recorded in almost every sheet except for two documents (Graphic 4.2). This information is recorded in the matrix diagrams which is ne the most codified pieces of information (Figure 4.1). After checking the presence of this data, no further qualitative analysis was made to check the validity of stratigraphic relations, even-though some contradictory relations were identified. This analysis was avoided because this topic has been recurrently addressed and there is already a well-defined strategy to deal with these problems (Roskams 2001; Harris 1989; Harris,Brown and Brown 1993).

Information about layer contacts is less frequent and includes descriptions of overlying and cutting relations. Nevertheless, the format only includes one predetermined field to indicate if a cut is truncated (10. Truncated, if known) (Figure 4.3). Some sheets follow this standard and they report when a cut has been: "*truncated to the south by modern concrete*" (TM-84) or "*truncated by [967]*" (SC-974). But a lot of them also describe when a cut truncates another cut or a layer: "*cut into brick earth surface*" (HK-256) or "*it cuts fill 251 and part of 252*" (SC-245). Moreover, many of these sheets also include graphic representations of cutting relations (Figure 4.13).

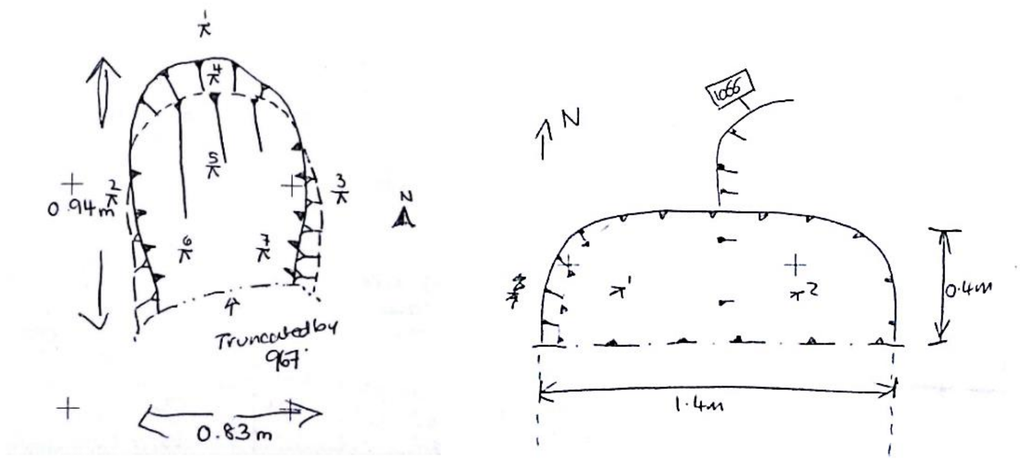


Figure 4.13. Sketches depicting cutting relation. SC-974 (left), ME-1119 (right).

Less often, sheets describe overlying relations among contexts: "*extends over bottom of cut to depth of 5-6 cm*" (MLC-147). Overall, the context sheet puts less emphasis in recording layer contacts, unlike stratigraphic relations that must be systematically recorded. Besides, the presence of layer contacts in every sheet is unnecessary because one document can summarize the physical relationships of two or more layers. Thereby, the absence of physical relationships (even in cut sheets) was not considered a sign of bad practice.

The sample included some examples of reflexivity about layer contacts (Graphic 4.4). These included a few descriptions of interpretative doubts due limited observational conditions as in the following example: "[truncated] *to the north and east by modern concrete, possibly by [1760], to south relationship uncertain*" (JA-1759) (Figure 4.9). Besides, none of these cases restrained the interpretation of stratigraphic relations. Additionally, there were very few examples when puzzling evidence produced uncertainty of layer contacts affecting other areas of interpretation like feature definition (Example 4.9).

The sample included more recurrent examples of uncertainty in stratigraphic relations. These cases were identified when stratigraphic diagrams included a question mark next to one of the relations (Figure 4.14). This representation rule is defined within the guidelines of the system (Museum of London 1994, section 3.1.1). Hence, these cases were classified as instances of evidential reflexivity because these doubts affect the credibility of the sequence. This was confirmed with two sheets that include additional notes explaining how puzzling evidence produced such interpretative doubt: "*up right supporting post for plank revetment [1131]. No cut* *is visible, seems to be hammered into the ground, therefore is difficult to say when this post occurred within the matrix*" (SH-1132). In contrast, the cause of doubt was more difficult to follow in sheets with no additional comments. Therefore, sheets that only included a question mark represent testimonies of evidential reflexivity but inadequately recorded (Graphic 4.4).

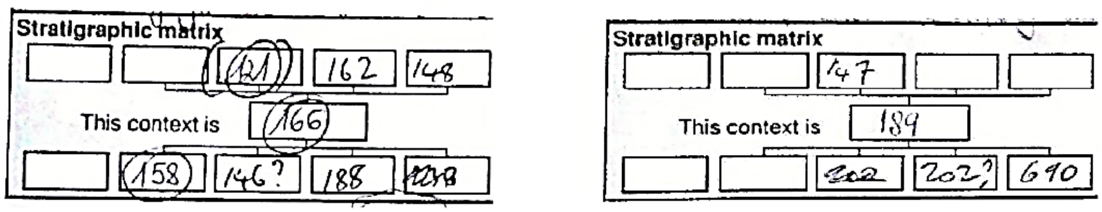


Figure 4.14. Examples of uncertainty claims in the stratigraphic sequence.

Nevertheless, the largest number of reflexive instances was coded from indirect testimonies or the corrections in the stratigraphic matrix. Most of these corrections were made by staff onsite, I know this from handwriting styles which belong to fieldworkers and supervisors. All these cases were classified as instances of trivial reflexivity, as corrections seemed to be a direct consequence of the process of discovery which produces reconsiderations in layer relationships and stratigraphic relations subsequently. This was suggested by the consistency in some related sheets, then when one relation was modified in one document, it was corrected in another too (Figure 4.15). Likewise, some sheets included additional evidence to support this idea. For instance, one sheet includes an edited textual comment regarding layer contacts: "*One of a series cut into layer ~~[1724]~~, [1666]*" (LAB-1702) and a subsequent modification in the diagram (Figure 4.16). No instances of retrospective reconsideration of stratigraphic relations were found; which in theory shouldn't be uncommon. Yet, it's possible that some of these situations are obscured among the sequence corrections but they are impossible to detect due the lack of a reflexive discussion.

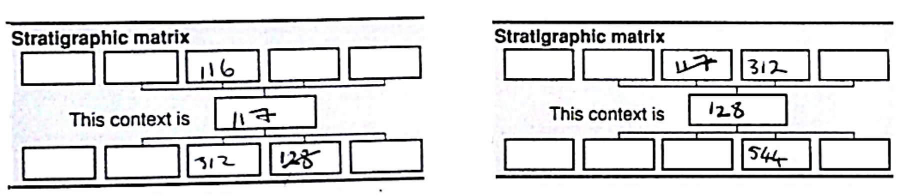


Figure 4.15. Corrections in the sequence over related sheets.

In short, single context recording prioritizes the description of stratigraphic relations over layer contact. Likewise, the system defines recording rules for ambiguous stratigraphic relations but a question mark without an explanatory account is insufficient for this task. Additionally, context sheets include numerous corrections over the sequence that represent indirect testimonies of observational fluidity. In general, these testimonies don't seem to require any more detail; except for possible cases of retrospective reconsiderations that possibly deserve an additional explanatory account.

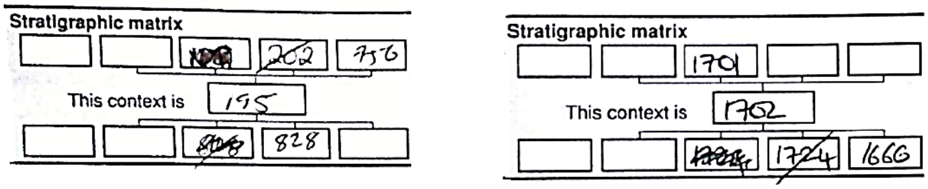


Figure 4.16. Corrections over the sequence.

4.5.6 Correlations

The percentage of correlations is very low (Graphic 4.2). These claims are symbolically recorded with the formula (x=y), either in the section 'context same as' or over the matrix diagram. Nevertheless, at least one of the related sheets always included a sketch to illustrate the correlation (Figure 4.17) and/or short notes to explain them: "*two areas flat metalled flooring both contexted as [321] but perhaps separate*" (PGF-144). Alternatively, some sheets correlated two contexts that belonged to the same layer but had been excavated in different interventions. For instance, a fill of a kiln "*recorded before fully excavated, same as 932*" (MW-852). Likewise, this alternative way to correlate contexts could be graphically recorded too as with [1115=1117], then making physical continuity evident (Figure 4.17). Both forms are considered within the guidelines of the system (Museum of London 1994, section 3.1.1). But more important, sheets always provided relevant background information to differentiate them. Hence, these testimonies were classified as cases of valid interpretations. There was a number of sheets that couldn't be assessed because only one of the pair was sampled (Graphic 4.3).

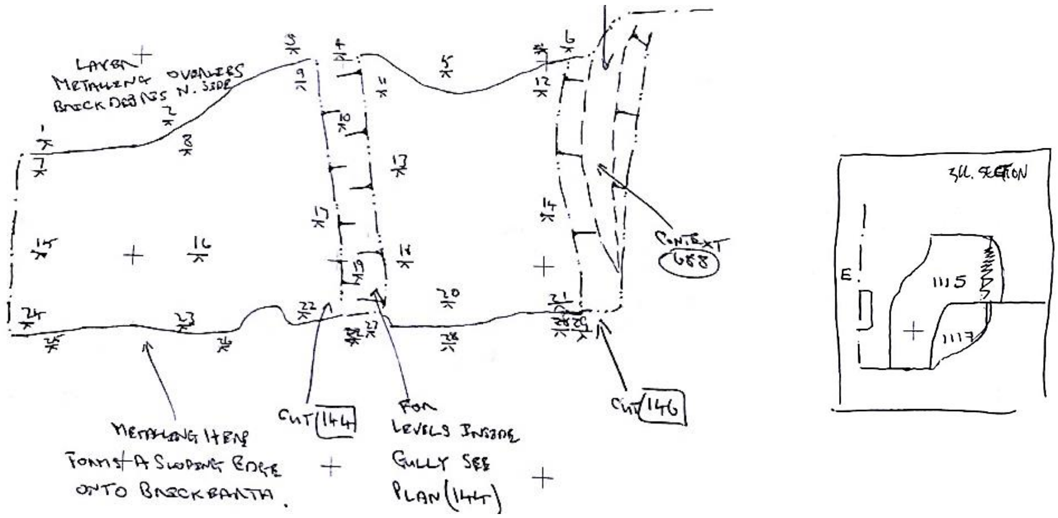


Figure 4.17. Sketches depicting correlated contexts. MLC-321 (left), SC-1115 (right).

Not many examples of reflexivity were recorded, but every example was related to cases of discontinuous layers. In general, uncertain correlations were identified with a question mark next to the symbolic representation (x=y?). Surprisingly, every one of these sheets also included an additional comment discussing puzzling evidence as in the example below. Hence, these cases were classified as evidential reflexivity (Graphic 4.4).

<u>Example 4.2</u>

> *"Possibly decayed timber/wattle lining of ditch. Much decayed matter within this matrix -probably originally structural. A more substantial possible lining is [530]. The patchiness of this fill and [530] (great variations in thickness) suggests they may both be parts of the same lining" (FAP/JA-523).*

One example of trivial reflexivity was coded from an updated sketch. Initially, a small brick earth layer [1010] was recorded by an excavator (PGF), but later another digger (MLC) discovered a larger portion of the same context interrupted by cut [189]. Then, he updated the description and included a new sketch for the correlated elements (Figure 4.18). In other words, this is an indirect testimony of fluidity along the process of discovery, but the example is even more important for illustrating the collaborative dimension of field practice.

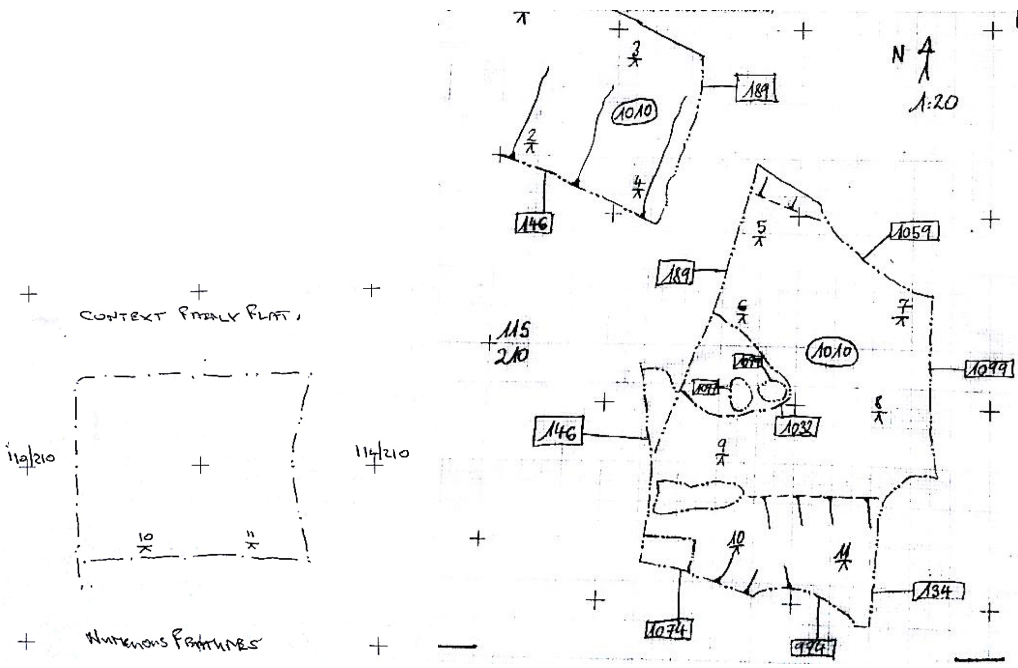


Figure 4.18. A preliminary sketch for layer [1010] (left) and the updated version (right). (PGF/MLC-1010).

4.5.7 Structural elements

Around 40% of the sheets describe structural layers as walls, floors, and timber linings and some of these layers form part of structural features like kilns, wells or postholes (Graphic 4.2). In general, a contexts sheet describes a structural component and two or more sheets can describe a structural feature. However, there is an important difference when reporting structures depending on the preservation of remains. When the type of structure is evident, context sheets include a description of its attributes and a short ‘interpretation’ of its type like a “*timber lining of pit*” (SC-124) (Figure 4.2) or a “*vaulted floor of kiln*” (HK-965) (Figure 4.19). These sheets normally include a sketch of the structural remain representing the thing being observed. Therefore, these sheets never include an interpretative conjecture, because they don’t need to explain what is evident to the sight.

Figure 4.19. The vaulted floor from the red kiln (S3). (Copyright MOLA).

On the contrary, when the structural character of a layer is not evident, sheets include an interpretive conjecture to explain how such remains indicate structural evidence. Frequently, these interpretations are based in marginal evidence like the compaction of layers or the distribution of inclusions like in the example below.

Example 4.3

> *"A bright red burnt layer suggesting a possible surface… Also, not complete due concrete foundation truncating around edge. Does seem to be a rough surface much slag in western half and pot all over suggest this was an attempt to consolidate the area, especially the western half" (FC-163).*

Some of these judgments use additional types of evidence like stratigraphic data. For instance, the example below considers layer attributes and the position of one layer in relation to another to interpret its structural character:

> *"Consolidated surface of burnt clay, pottery and gravel. Uniform surface layer of industrial debris (burnt furnace/kiln clay and some gravel. Possible foundation for repair metalling above [194]. Appears like layer of path-type structure leading towards the kiln" (MLC-603).*

The interpretation of structural cuts was also a recurrent activity at MRG95. For instance, the example below interprets cut's attributes when no timber-post had survived and packing evidence was marginal (Figure 4.20).

Example 4.4

> *"Very substantial posthole, most likely for a structural purpose since size and depth suggest a weight bearing purpose. Layer of cobbles at the base for support. Not found elsewhere at that level" (LAB-1656).*

Likewise, when layer evidence was very fragmentary, this could trigger alternative interpretative processes leading to consider less typical signs. For instance, the testimony below explains the interpretation of spatial signs alongside layer attributes for defining the structural character of a decayed fragment of timber.

Example 4.5

> *"Evidence of decayed planking and surviving timbers (1131+1132+1133 plus others not yet numbered) suggest a timber structure existed, possibly part of a revetment quay. Adjacent to the watercourse (Walbrook) nearby" (FAP-1118).*

Unlike B49 that mainly included evident structures, MRG95 included a mix of evident and marginal structures. Hence context sheets only describe well preserved structures but when evidence is marginal or fragmentary, context sheets must have an explanatory account. In the last case, sheets included a wide range of signs to support interpretations, many of them scaping the 'rigidness of the format.' Another surprising aspect is the low proportion of unreliable structural data. This indicates that fieldworkers are very careful to record this type of remains (Graphic 4.3). The few problematic examples usually represent cases of non-evident structures in which documents fail to explain the interpretative process. For example, one sheet interprets a deposit as a 'makeup layer' (LD-1569) but no further information was offered besides the description (firm, dark greyish brown sandy clay silt layer). Some of these documents only encircled the 'structural' field in the interpretative section of the format.

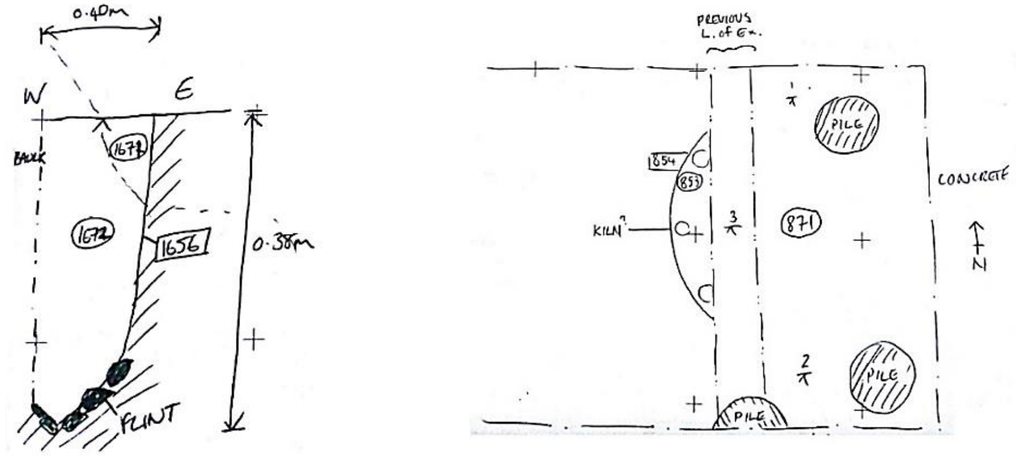


Figure 4.20. Sketch of a posthole. LAB-1656 (left) and sketch of partially discovered kiln. MW-871 (right).

Context sheets also include reflexive testimonies about structural elements. One of the most distinctive examples of observational reflexivity was a preliminary sketch that shows a partially exposed feature (S3) and includes a short annotation claiming uncertainty about the type of structure being excavated: “kiln?” (Figure 4.21). Evidently, this doubt became irrelevant after exposing the full element. More frequent are testimonies when fragmentary evidence produces doubts about the structural character of vestiges and therefore fieldworkers include a short discussion of evidence: “*Ditch lining, wattle hurdles possibly. No evidence of solid timbers, but soft decayed matrix which could be same as [523]*” (JAP/FAP-530). Specifically, this discussion seems incomplete, but this testimony must be read in relation to an additional reflexive discussion recorded in another sheet (see Example 4.2). Similarly, some sheets discuss puzzling evidence of cuts and deposits alongside (see fragment below and Example 4.7). Altogether, these cases were classified of examples of evidential reflexivity

> *“Pit-cut? very truncated so difficult to tell exactly what shape this feature may have been originally. A slightly rounded west end suggests some kind of pit. Evidence of whittle and stake hole suggest a lining existed” (FAP-820).*

There was an atypical example with a bunch of related textual and graphic testimonies describe the interpretive steps in the definition of a posthole. In some extent, these testimonies reflect the act of discovery (Figure 4.21. A series of related sketches depicting various contexts of a posthole. An excavated cut with the mark of a stone slab (left). The upper surface of a fill of stones overlying the cut (centre) and a squared slab in the base of the cut (right). JA-1421 (left) and JA-1437 (centre and right)Figure 4.21). But the last textual testimony (JA-1337b) makes a reconsideration of the structural character of the stony fill (Example 4.6). Hence, this testimony represent an reflexive account to contextualize the credibility of an ambiguous interpretation based in fragmentary evidence, although the interpretative reconsideration is not explained very well.

<u>Example 4.6</u>

> *"Possible posthole. [1421] might have been cut for large stake. [1437] fill had appearance of packing for a large stake" (JA-1421).*
>
> *"Fill of [1421] Contains several blocks of building material CBM. Possible posthole with packing material" (JA-1437a).*
>
> *"Diagram over shows slab founds at base of [1437]. This slab was beneath soil as opposed to cobbles. Possibly this was slab at base of stake, thus no packing material above. Slabs sometimes used to rest wood posts on stop damp attacking wood" (JA-1437b)*

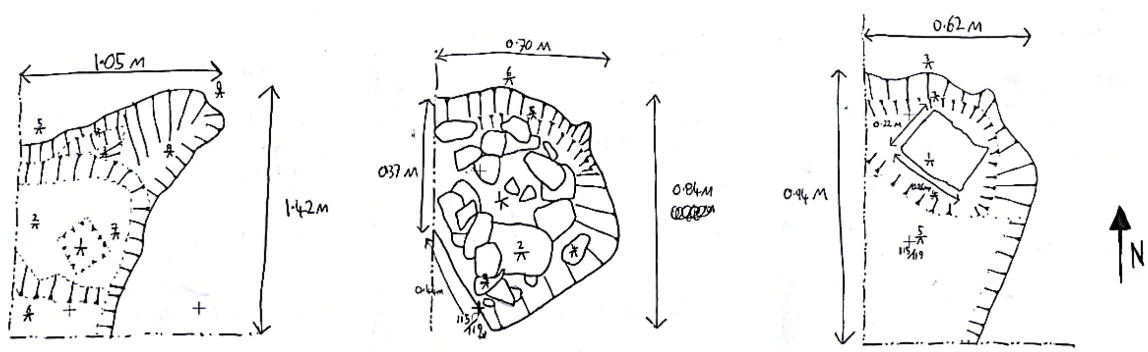


Figure 4.21. A series of related sketches depicting various contexts of a posthole. An excavated cut with the mark of a stone slab (left). The upper surface of a fill of stones overlying the cut (centre) and a squared slab in the base of the cut (right). JA-1421 (left) and JA-1437 (centre and right).

The most recurrent problem with reflexive testimonies is when an ambiguous situation is reported but unexplained (Graphic 4.4). These was specially recurrent in some cases when the strcutural character of a remain was evident but its type was uncertain: "*possible* [burning] *pit or termination flue*" (SC-1372) or a wall probably defining a *"cellar wall/cess pit*" (L'ON-1464) (Figure 4.22). Generally, these sheets include sketches illustrating the structure but they lack a discussion of such interpretive possibilities; therefore, being classified as cases of problematic reflexive record (Graphic 4.4).

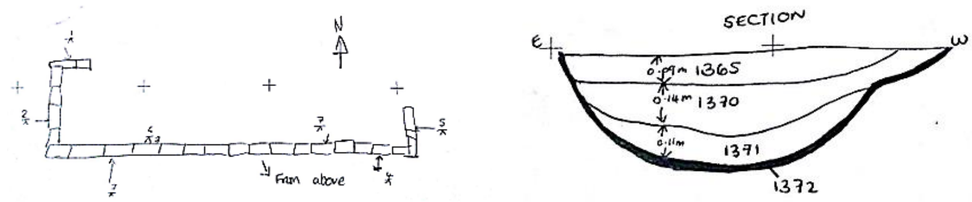


Figure 4.22. Sketches of structural remains. L'ON-1464 (left), SC-1372 (right).

4.5.8 Features

Feature information covers around 40% of the sheets (Graphic 4.2). This information is recorded using different procedures, but formats put more emphasis in recording group relations for negative features. Specifically, the cut sheet includes a section within the descriptive box to list its associated fills. For instance: "*835, 895, 888, 877, 875, 857*" (LMC-904). However, it's frequent that fill sheets repeat grouping relations as with "*primary fill of [904]*" (LMC-835) and the "*penultimate fill of pit [904]*" (LMC-857). Nevertheless, context sheets almost systematically describe group relations for positive features like: "*lower infill of kiln [1026]*" (MW-932), the "*floor of kiln [1026]*" (LAB-1214) and so forth. Then, by following these links, one can group contexts from a feature. Likewise, by following feature relations and stratigraphic relations, one can derive the physical relationships of components generally. However, some sheets facilitate this work by including a sketch of the feature (Figure 4.23). This demonstrates an intrinsic flexibility in the system for recording higher order entities by introducing different form of contextual relations among units.

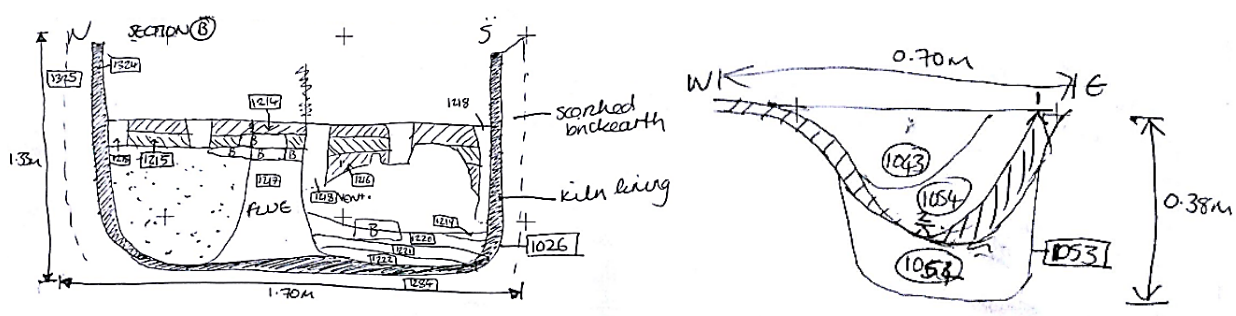


Figure 4.23. Sketches of features. LAB-1026 (left) and LAB-1054 (right).

The documental analysis was crucial to understand that feature interpretation is more than defining a group of contexts, because records often mention the type of feature being defined.

These can be divided in two large classes: structural features (kiln, well, hearth and so forth) and non-structural features (burial, rubbish pit, midden and so forth). In general, MRG95 was a rich source of negative features and cuts provided important evidence for interpreting feature-types, especially when layer's evidence was marginal or inexistent. In general, structural features were interpreted from cuts with sharp, vertical and stepped edges.

Example 4.7

> *Although it was difficult to determine exactly what this cut was intended for, its general shape indicates that it could have been a posthole. It is however particularly deep considering that there are no indications of any major structural features" (AH-1000).*

On the contrary, negative features with irregular cuts commonly were associated with non-structural features. Nevertheless, an interpretation considering evidence form distinct layers was more solid than considering the evidence of cuts alone (Figure 4.24).

Example 4.8

> *"This looks like a half circular rubbish pit. Filled with domestic and industrial refuse. It is too large and irregular to be a terminus or structural cut and contained lots of dry rubbish fill" (MW-1452).*

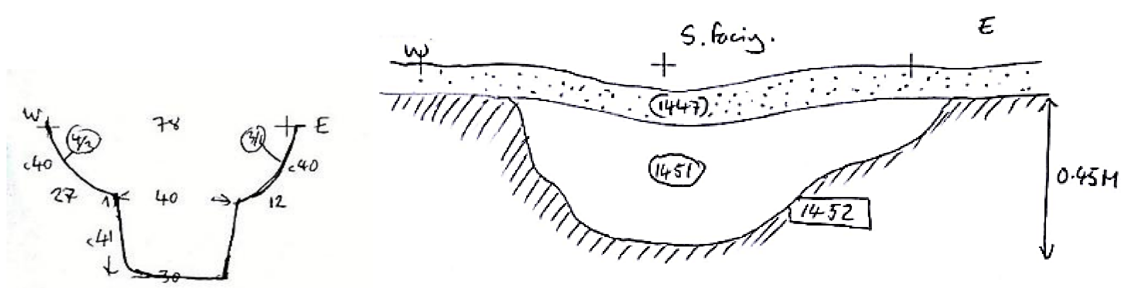


Figure 4.24. Cut of a structural feature. AH-1000 (left) and sketch of a non-structural feature. MW-1452 (right).

Feature interpretation is one of the clearest examples of fragmentation of the interpretative process into a series of related sheets. However, it's common that cut sheets provide a synthetic explanation of the feature-type (Example 4.8). However, when a cut sheet provides an unexplained interpretation, one must examine the information in related documents to evaluate the credibility of that interpretation. For example, cut sheet [97] mentions describes the type of a feature, but the explanation is recorded in fill sheet [76].

> *"Contains a lot of leather, also bone and pot. Within layer, light brown flecking suggesting an eroded edging to pit, although possibly slumped with sides being irregular and a lot of machining acting in the area…Pit fill of 97. Some type of leather dump" (FC-76).*

> *"Cut of pit. Probable of leather waste pit" (FC-97)*

Generally, the main problem with feature claims is when the class or type of feature is not considered in any of the grouped sheets. This problem was recurrently observed for negative features that often are described as pits or ditches. However, this kind of nomenclature is inadequate for not being indicative of the structural or non-structural charter of features. In many of these cases, all the related sheets use a similar terminology as in the following example: "*fill of pit [345]*" (RC-344) and "*cut of a small shallow pit*" (RC-345). These cases represent typical examples that critics use to argue that standardized sheets restrain field interpretation to claims like: "this is a pit" (Lucas 2001, 8). However, only one fifth of the feature claims fits with this pattern. This type of problem was inexistent for masonry features which commonly indicate its type, and this suggests a harder challenge to investigate non-structural features. Finally, there was small portion of sheets that couldn't be examined because only one sheet of a group was sampled (Graphic 4.3).

Cases of feature reflexivity describe different interpretative challenges produced by puzzling and fragmentary evidence (Graphic 4.4). Some cases describe when the class or type of feature was uncertain. For instance, the fragment below briefly discusses evidence for a non-structural pit, but the situation is equivalent to uncertainty for structural features.

> *"Pit. Uncertain purpose -possibly dug for brick earth although unlikely since naturally not encountered within this trench. Fill [1269] very peaty" (LAB-1357).*

Unfortunately, many sheet describing similar situations only reported an interpretive doubt without further discussion: "*only fill of feature [226], possibly a small pit or a large posthole. It is heavily truncated*" (RC-225). Hence, these examples were counted as instances of problematic reflexive recording. Another interpretative challenge was when the number of features was unclear, quite often due ambiguity in layer definition or layer contacts: "*Pit, uncertain function, possibly structural. Could be a part of cut [197] but seems to be a separate feature*" (HK-758). In such cases supplementary sketches resulted very helpful to explain reflexive dilemmas.

However, these testimonies were not always successful to discuss positive or contradictory evidence, or this was minimally considered (Figure 4.25). The example below is one of the best examples when a sketch and textual testimony work together to explain an interpretive doubt and discuss evidence.

<u>Example 4.9</u>

“Two interlinked stake holes. No obvious relationships to each other. Fills very similar, size similar. Therefore, suggested these operated together as a stake hole” (PGF/MLC-804).

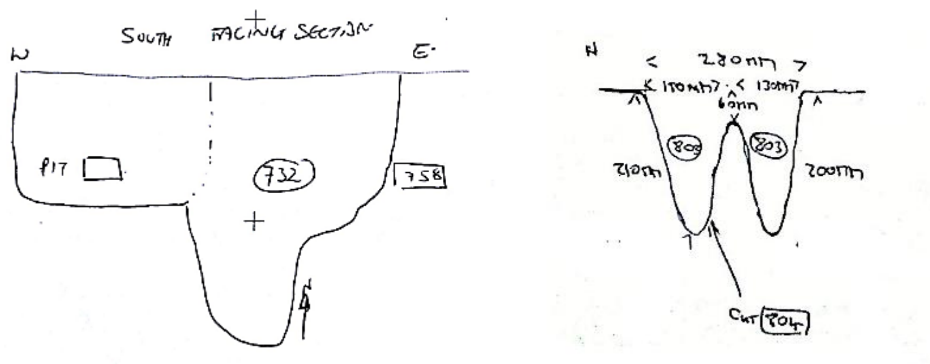


Figure 4.25. Sketches depicting uncertain situations of feature definition. HK-758 (left), PGF/MLC-804 (right).

Finally, another form of reflexive challenge arose when it was unclear whether a layer formed part of a feature. Often, these issues where textually explained: “*Slump in the side of ditch [793], possibly part of ditch [546]. But recorded as separate feature*” (LW-760). Ocassionally, an explanation was complemented with a sketch illustrating more clearly that fill [1055] was not completely contained by cut [832]: “*After examining section 20, it appears that this could be a subsided layer or fill of [832] into clay lined feature 1056/1213/1214. Hence revised matrix*” (MW-1055) (Figure 4.26). As in former situations, some reflexive testimonies are better explained than others. Then, this is another area where the quality of reflexive testimonies is inadequate revealing a huge number of recording problems (Graphic 4.4). Independently of this, MRG95 showed a wide diversity of interpretive and reflexive situations due a wide diversity of explored remains. For example, many of these reflexive challenges were practically absent at Catalhöyük where negative features were commonly sealed.

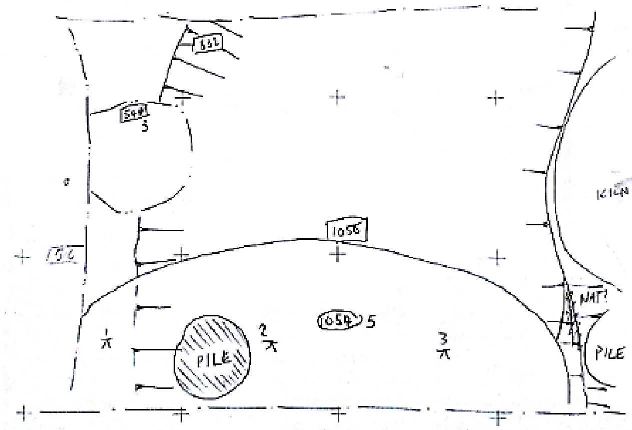


Figure 4.26. A sketch representing an ambiguous feature relation between a layer and a cut (MW-1055).

4.5.9 Formation processes

Formation processes claims covers 50% of the sample (Graphic 4.2). However, this number should be read considering the following elements. In general, cut sheets don't include thin information except for natural events like holes produced by tree roots. Nevertheless, there is a systematic investigation of the structural character of cuts. Similarly, records of structural layers usually do not include this information but none of these were considered problematic because claiming the structural character of something is an implicit description of its formation process. Nevertheless, cut sheets described formation processes occasionally. For example, to report an atypical constructive process: "*flue appears to be tunnelled as there isn't clear cut visible after removing furnace lining*" (MW-545), but more often cut sheets describe post depositional transformations: "*the SE edge is fairly straight whilst the SW edge is sloped. It is possible this is due to the infill and disturbance*" (FC-97).

Likewise, structural sheets occasionally included depositional information for describing evidence of use events, for instance, the "*dirty grey nature of surface suggests use of metalled surface over time. Further corroborates observations that density of metaling may increase towards east and south extend of the trench*" (MLC-856). Modifications and reparations of structures were also occasionally considered: "*small patch of metaling located to the south and above metalled floor [321]. Not cut by any feature. Yet it finishes on all sides*" (PGF-329). And finally, destruction events as a "*collapsed kiln roof. Upper part of the kiln had collapsed down onto the performed floor [965]*" (HK-903).

Overall, most formation process data emerged from the investigation of sedimentary layers and the first step consisted in defining the natural/cultural character of deposits. When testimonies described natural layers, sheets provided minimum evidence to subscribe such claims, for example a "*natural clay deposit with occupation debris. Part of the horizontal silting layers adjacent to the watercourse of the east*" (FAP-755). On the contrary, cultural layers usually required more analysis to define its primary or secondary deposition. Secondary dumps were among the commonest form of cultural layers, and its depositional process was largely inferred from layer attributes and finds content.

> *"Fill of 1708 which has possibly a slumped into a cut. A characteristic dump containing dump material from a furnace, furnace wall lining in the form of burnt clay attached to vitrified glass" (SC-1736).*

In general, the interpretation of formation processes always was grounded in signs from visual and tactile experience, but a few examples included less common evidence derived from olfactory experience like the following example: *"waterlain deposits indicated by stickiness, wetness and general bad smell*" (FAP-915). Overall, former examples show very clearly the interpretative possibilities of embodied experience. However, the interpretation of sedimentary layers provided very clear examples in which the skill of fieldworkers to interpret marginal evidence is very clear.

<u>Example 4.10</u>

> *"Very striking mixed deposit of three industrial deposits dumped in one location, possibly representing waste components of the same process. Clayey silt components appear to be buried turf used on fire, since comparatively like charcoal. Very humic material, peat from marsh or fuel. Semi-rotted ash from hearth" (MLC-190).*

Previous examples also demonstrate that formation process claims are derived from layer, finds and stratigraphic data and given the diversity of signs and depositional process, there is no mechanic way to interpret them. In this way, reliable testimonies always included an explanatory account linking evidence with an unobservable depositional process. On the contrary, the commonest form of invalid data are unjustified depositional interpretations (Graphic 4.3). In such cases, sheets include a short interpretive statement like "*primary fill of small pit*" (1726-LAB) without proving an explanation. Additionally, there were some unreliable sheets of sediment layers but that only described post depositional transformations without considering formation processes, for instance a "*top fill of wooden lined pit [81] which has been truncated by modern foundations*" (RC-82). Thereby, unreliable depositional data was frequently related to sedimentary layers which represent more complex interpretive challenges and receive less systematic attention that structural evidence.

At this point, one can notice the close relation between invalid layers descriptions, invalid structural claims and invalid formation processes claims. In specific, invalid layer descriptions represent an absence of depositional interpretations, whereas invalid structural claims are unjustified claims of a specific depositional process and finally, unreliable formation process data are failures to demonstrate the primary or secondary origin of a layer. On the contrary, a valid layer description is the first necessary condition for a valid structural or formation process claim. Then, by summarizing the number of valid layer descriptions, strcutural data and depositional claims in one group and the number of unreliable testimonies in another, it was possible to get a general picture of the credibility of testimonies in basic areas of interpretation which is defined a General Index (Graphic 4.3). The general index shows that two thirds of the sample are reliable testimonies. The analysis of records also identified additional forms of unreliable testimonies like invalid feature descriptions, invalid functional and agency claims. However, these do not form part of the general index because these types of mistakes are generally derived from mistakes in basic forms of interpretation (layer description, structural and depositional data).

Finally, some sheets included reflexive testimonies about formation processes and many of these examples describe interpretive doubts produced by fragmentary and puzzling evidence (Graphic 4.4). For instance, when it was unclear whether some remains were archaeological: "*a possible irregular feature. the fill contained a lot of modern disturbance. It is not clear whether it is modern or archaeology*" (RC-86), or when it was ambiguous if some remains ere cultural or naturally deposited: "*Probably tree/root disturbance from medieval garden phase although could possibly be a cluster of stakeholes. Most probably actually!*" (LMC-224). However, the most interesting cases examples of reflexivity included discussion of interpretive possibilities because evidence was ambiguous.

"Possibly construction packing/debris behind ditch lining. A lot of fragmentary inclusions in this layer suggesting debris falling behind decaying hurdles? Alternatively, this could be a purposeful 'packing placed' behind the lining, when erected. Would bits of animal bone and such be used for this? (FAP/JA-531).

Additionally, depositional information produced one of the most interesting examples of fluidity. Initially layer [242] was interpreted as a dump filling a hole. However, after removing the fill, the cut was interpreted as a worn-area, and the fill was reinterpreted as a patch rather than a dump. In other words, this example represents a retrospective revision of beliefs related in the light of new evidence.

Example 4.11

"Fill of 243. dumped industrial waste filling a hole /worn area in the brickearth" (HK-242).

"Very irregular cut-appears to be worn rather than deliberately cut. The fill [242] seemed to be patching a hole in the brickearth" (HK-243).

Reflexive content also included cases of problematic testimonies in which a relevant reflexive situation was reported but not discussed. For instance, with an "*external dump layer/occupational layer*" (JDM-747) or "*gravel surface or dump?*" (L'On-1627). Hence, these testimonies were classified as problematic instances of evidential reflexivity.

4.5.10 Groups of features, spaces, buildings and phases

A low proportion of sheets described groups of features (Graphic 4.2). In most cases, these defined groups of stakeholes and postholes based in its spatial proximity and/or physical similarity (Figure 4.27). In other words, many groups of features are based in spatial data and middle level contrasts Most of these claims are recorded with textual testimonies and sketches that altogether describe the rationale of the interpretative process. In general, groups of features defined structural elements such as a "*four stakeholes similar in nature […] all approximately 0.55/0.60 meters apart forming a square structure*" (LMC-401) or a group post-holes "*in line running north south, evenly spared*" (PGF-991).

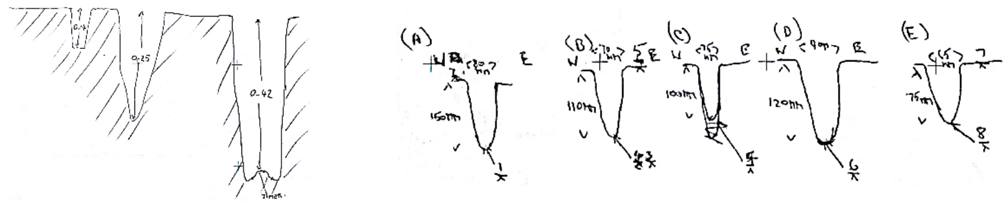


Figure 4.27. Sketches of groups of features. AH-988 (left), PFG-991(right).

Quite often, groups of stakeholes and postholes were recorded as single events and each group was described in one sheet reporting multiple cuts and fills. This strategy was based in the premise that fills and cuts were alike but when required, relevant differences were mentioned. Overall, this represent a violation of recording standards as each event should have been recorded on an independent sheet, but this economic approach didn’t produce affect the credibility of data. For example, all the associated cuts always cut over the same layer, therefore not compromising stratigraphic data. Instead, when stratigraphy was more complex, the traditional recording strategy was followed. Likewise, some of these sheets had complementary notes explaining this: “*one of a group of postholes in the NE corner of the trench, appear to be stratigraphically contemporary, although they differ in depth and character*” (LAB-1598). In general, most group of features claims seemed rather speculative as there was not a real clarity what types of structures this could have represented. Hence, this was the commonest form of reflexivity, but this was not very frequent (Graphic 4.4): “*Although it’s possible to hypothesise over various potential associations the evidence to form specific structures here is not very strong enough to irrevocably link individual stake holes*” (LMC-382). Nevertheless, the most interesting example is one testimony that hypothesizes a possible walkway. Overall, this is one the most sophisticated exercises of reflexivity and interpretation, based in fragmentary and marginal evidence recorded in brief discussions.

<u>Example 4.12</u>

> *“Posthole. No obvious structural remains…However, lots of peat found in this grid square (fills of water-filled pit), and these posts maybe part of a walkways across marshy ground. Area too busy to be certain of relationships but [1316] cuts flue fill and is cut by [1205] and probably [1354]” (LAB-1316).*

This discussion was complemented with a sketch illustrating the spatial patterning among two postholes accompanied with a note that indicates: "s*imilar post-located 1.16m to the SW"* (LAB-1316) and this note was very helpful to understand the sketch. Hence, both testimonies work together to explain the interpretation of a possible structure defined by a group of features and grounded in spatial and depositional evidence (Figure 4.28).

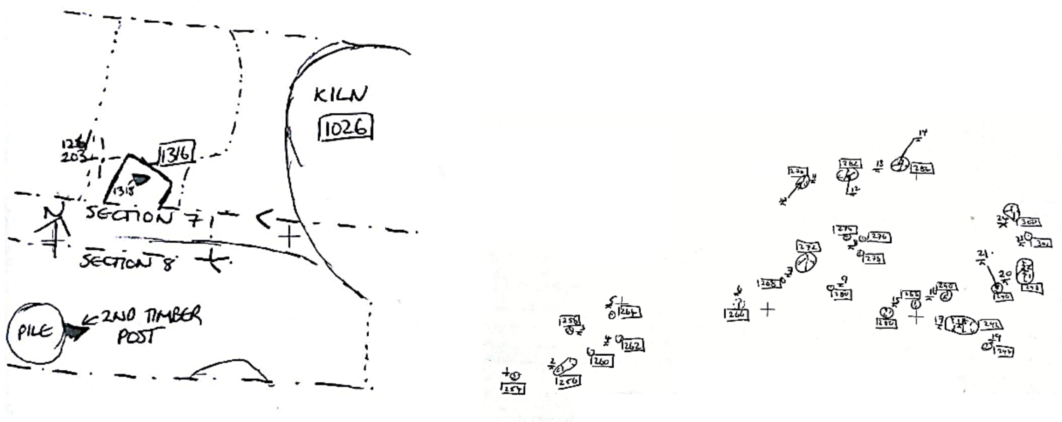


Figure 4.28. Sketches of groups of features. LAB-1316 (left), HK-254 (right).

As explained before, a group of features is an auxiliary concept to define structures, areas and spaces. Then, when groups don't define none of these, they can be difficult to understand. For instance, one sheet describes a "*shallow ill-defined pit…. Maybe associated with pits [727] and [729]" (LAB-1725)*, however none of the associated records explains whether this group represented an activity area, for example. Hence, this and similar examples were classified as instances of unreliable testimonies (Graphic 4.3). Nevertheless, one of these unexplained groups provided one of the clearest examples that even problematic records are not completely useless during post-excavation. Specifically, an un-patterned concentration of similar stakeholes cutting a brick surface (Figure 4.28) was the basis to define an activity area in post-excavation (Figure 4.5).

The brickearth surface was pierced by a large number of stakeholes which formed no discernible internal structures or pattern. These are thought to represent the potters working area possibly associate items of furniture such as tables and shelvings for stacking drying vessels (Seeley and Drummond-Murray 2005, 14).

Spatial information was reported in one third of the sample (Graphic 4.2). Generally, these examples only indicated whether a layer or a feature was internal or external according to the predefined categories in the sheet (Figure 4.1). Then, most of these sheets do not provide further explanation or evidence, hence being considered as unreliable testimonies (Graphic 4.3). This is one of the areas where one can see more clearly the negative effects of the codification of data. Yet, even such problematic data was not completely mute, for example many sheets describing rubbish dumps were classified as external vestiges. A similar situation was observed for spatial reflexive claims, because records generally include an additional question mark next to the predetermined fields to indicate uncertainty. Therefore, representing instances of recording problems for reflexive claims (Graphic 4.4).

The best testimonies of spatial information are a few sheets with short annotations describing specific aspects of the spatial organization of the site. For instance, one sheet provides an indirect mention of the main activity area of the site: "*forms central lump of redeposited brickearth to immediate W of industrial furnace complex*" (MLC-1010) (Figure 4.6). Likewise, another sheet roughly indicates the main internal activity areas defined by kiln structures and dumping pits: "*uppermost remaining fill of re-cut [522]. Appears to be a dump of pottery and glass presumably pot wasters from nearby kilns*" (LMC-241). Additionally, some sheets explain the access route to the furnace complex: "*appears like layer of path-type structure leading towards kiln structure to E*" (MLC-603). Whilst others describe the relation between some linear features with the watercourse like a "*possibly drain leading down to Walbrook to the NE*" (LJB-709). Finally, some sheets conjecture about the extension of the site considering the presence of structural remains in one side: "*area of early mid period metalling on edge of marginal part of the site, but having traces of industrial activity*" (PGF-688); and the limits imposed by the river on the other (Example 4.5). It's clear that much of this spatial information only make sense when being read in relation to background information provided by additional sheets. For instance, the reference to the furnace complex only makes sense if one is aware of the existence of kiln structures. Then, this is one of the aspects where context sheets do not seem to reflect the quality of spatial investigation onsite; and possible this is one of the aspects that could have benefited from an additional recording tool like a notebook, especially in a site where spatial organization is not clearly defined by architecture.

One of the clearest differences between field and post excavation phases at MRG95 is the interpretation of an open tiled roofed building offsite. This building, arguably a potters' workshop (S1), was defined by two rows of postholes, one vertical and one horizontal around a brick earth surface (Figure 4.5). Thereby, building data is an important difference between B49 and MRG95. However, B49 was evident since the earliest stages of excavation, whereas perceiving a 'potters' workshop' onsite was more difficult due the fragmentariness of remains. Nevertheless, this produced a strange similarity in the archives of both projects because B49's diaries do not discuss building evidence due its clarity whereas this possibility escapes MRG95 due its difficulty. In short, these examples clearly show how the character of remains plays an important role in the interpretive process of a site. Hence, the absence of building information at MRG95 cannot be necessarily understood as a sign of bad practice (Graphic 4.2.

Phase information is another area where there is an important contrast between site and offsite investigation at MRG95. Specifically, this type of claims is practically absent from site records; whereas the published report defines multiple constructive phases in the pottery workshop. For instance, an initial roman phase of rounded kilns (Figure 4.5) followed by another of 'bottle shaped' kilns (Figure 4.29). However, these limitation also seems to be explained by the character of site which is more complex than B49.

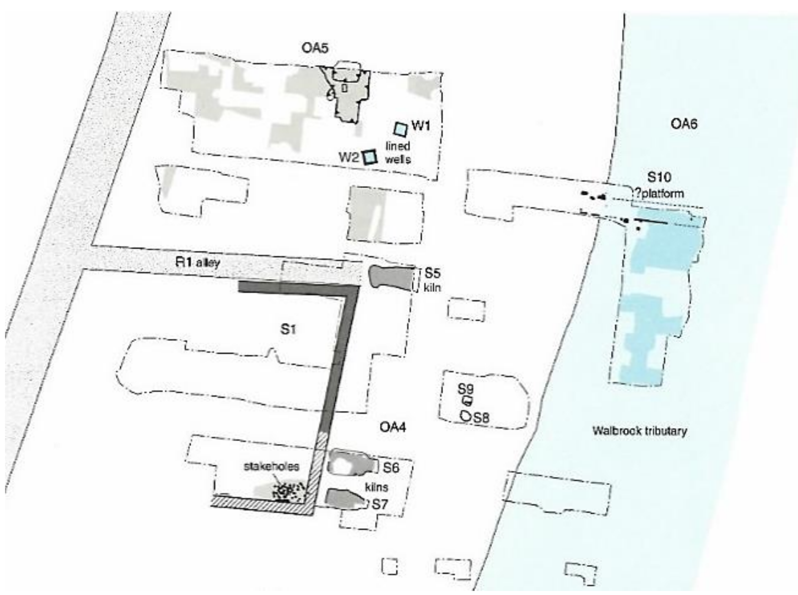


Figure 4.29. MRG95. Roman Period. Phase 2 (MOLA copyright).

4.5.11 High-level claims

High-level data includes functional, agency and culture/period claims, but the frequency of all of them is very low (Graphic 4.2). In general, these claims are recorded in the interpretative sections of the sheet, but the standards of recording variated. Culture/period claims describe cases when finds and remains are interpreted in relation to a cultural group or a historic period (prehistoric, roman, medieval). In general, sheets only include short comments as "*post-medieval brick well*" (LD-1455), "*fill of flue (Medieval)*" (SC-1330), "*floor joints of roman kiln*" (L'On-1020), without providing any explanation for this. Likewise, culture/period reflexivity is recorded with short notes suggesting a tentative possibility "*Roman gully?*" (LD-913), a "*Prehistoric Layer? Not sure*" (LAB-1755) or a threshold of options: "*Dump layer Medieval/post-Medieval*" (HLB-1345). However, these were not classified as valid/invalid data because this type of data is the domain of post-excavation domain in this methodology. Otherwise, if a higher standard were imposed (as with reflexive method), then every culture/period claim should be considered unreliable (see section 3.5.10).

In contrast, functional and agency claims were commonly recorded with an explanatory inference supported from different kinds of evidence. Agency claims included a few instances of structural and depositional evidence reinterpreted in terms of human action. For instance, "*a dump associated with kiln area consisting of rubbish (breakages/wasters) from cleaning out kilns*" (MW-323) and some tool marks in stakes and posts were interpreted as traces of manufacturing actions (Figure 4.12). Functional data is the most frequent type of high-level information. However, site records usually incline to investigate the function of non-structural features. One of the commonest functional data was the interpretation of rubbish pit, finds and depositional evidence. However, these vestiges also provided indirect evidence of other industrial activities in the site: "*The glass and vitrified glass and crucible suggest possible glass production in a nearby vicinity. This is a roman dump layer*" (SC-1405). Equally interesting are the examples of functional interpretation based in marginal evidence of sedimentary layers.

<u>Example 4.13</u>

> *"Very deep quite regular steep sided cut. Appears to be clay lined with thick silty clay deposits in base. Higher peaty fills suggest water management possibly associated with waterlogged post [1318]. Storage/setting pit for raw clay to make pots?" (MW-1056).*

Overall, valid functional testimonies always include an explanatory conjecture whereas invalid functional data includes cases when functional claims are unjustified (Graphic 4.3). One consequence of the functional imbalance between non-structural and structural vestiges is that almost every unreliable functional claim is related to non-structural elements as in the following examples: "*fill of a rubbish pit*" (TM-83) or a "*domestic rubbish pit*" (HLB-519). Many of this unreliable functional information derives from problematic layer descriptions and depositional data.

The functional imbalance is also reflected in reflexive claims and sheets record more often interpretive doubt and ambiguities for non-structural features. For example, the "*function of pit unclear. Possibly quarrying for brick earth, in fact dug through redeposited brick earth and a layer of gravel and backfilled with gravel*" (PGF-146). Another interesting example was the reflexive testimony of a rectangular pit with fills of dirty grey clay: "*Primary fill of pit [904]. Although this may have been a lining, the subsequent truncations suggest that this is more likely to be the remains of a clay stored pit prior to use*" (LMC-835) (Figure 4.6).

Almost one half of reflexive functional testimonies were poorly recorded, particularly when functional ambiguity was acknowledged but not discussed (Graphic 4.4) like in the following examples: "*a pit was dug for the purpose of a possible rubbish pit or as quarrying pit to obtain brick earth and subsequently filled with rubbish*" (SC/JA-1760) or a "*massive Ditch cut. Unclear as to ditch's purpose, may be defensive. Possible drainage into Walbrook tributary*" (HLB-546). Yet, the commonest problematic statement is when sheets report functional ignorance without the slightest explanation: "*Pit cut. Function not known*" (HLB-978) or a "*small feature of unknown function*" (LJB-521). These problem with these testimonies is failing to explain such lack of knowledge, then being unclear whether these were situations of puzzling evidence or a fieldworker with insufficient interpretative skills. This distinction was acknowledged in one atypical testimony that acknowledges the limited background knowledge produces doubt (see below). This testimony is valuable because it provides a fairer description of an epistemic situation. Obviously, having site records that could make this distinction would be desirable, yet whether this is feasible is difficult to say, because epistemic agents commonly are not aware of their lack of interpretive skills in the first place.

> *"Rectangular pit, possibly created for storage of "raw" clay, may have been used for puddling although a limited knowledge of the subject dashes me speculating much on this one –I would have expected something dirtier though" (LMC-968).*

One of the differences of SCR with the feature system is the lack of an additional format to describe features and their depositional history (Figure 5.2). However, if individual events are properly investigated, SCR can be quite successful to record these aspects in a series of related sheets. Generally, context sheets provide nice testimonies of the depositional history of masonry features because fieldworkers commonly make neat distinctions between constructive, refurbishments and destruction events. However, when the type and function of a feature changes over time, these aspects can be nicely recorded too like in the example below that describes a structural feature that later was reused as dumping rubbish.

<u>Example 4.14</u>

> *"Fill of wooden lined pit, this is a characteristically industrial fill, plenty of animal bone and ovicaprid horn in abundance, possible waste pit? No nails or scraps of leather found so doubtfully for tanning. Possibly suggest near butchery or food preparation" (SC-123). (Figure 4.1).*

> *"Timber lining of pit" (SC-124) (Figure 4.2)*

> *"Cess pit. Originally excavated and recorded as timber lined pit. Fill [794] removed much later. Pit was cut through dump layers into natural brickearth which was stained green" (HK-134) (Figure 4.3).*

In sum, although the industrial character of the site had been defined during evaluation work (Lakin 1995), the excavation of MRG95 was paramount to investigate the diversity of industrial activities of the site. Specifically, this was achieved by the investigation of features and structures that altogether helped to interpret the general function of the site as a pottery workshop based in the evidence of storage, production and discard activities. Thereby, the project shows an important change in its epistemic status before and after its excavation. Besides, a lot of field data became cornerstones for final interpretation of the site in the published report: "*The area in front of Structure 1, Open Area 4, had a large number of pits and dumped deposits, indicating that this was the main area used to store and levigate (settle-out) clay and their respecting waste products*" (Seeley and Drummond-Murray 2005, 26) (Example 4.13; Figure 4.6).

4.6 The importance of documental analysis.

The purpose of this chapter was to evaluate two claims about standardized recording methods and more specifically about single context recording. The first claim is the supposed impossibility to assess the credibility of primary data in context sheets. To demonstrate that it is possible an intensive analysis of context sheets was implemented based in a confrontation of descriptive and interpretative data. This procedure is inspired in the general methodology espoused by history and jurisprudence based in the cross-examination of testimonies to identify consistencies, errors and falsehoods (Ginzburg 1999, 12–13; Bloch 1954, 51). This procedure allowed to differentiate between valid and invalid testimonies. Besides, the analysis allowed to identify three standards for reporting primary data. In some cases, a descriptive testimony of the thing being observed is sufficient, for example when reporting well-preserved structures, however when remains are more fragmentary an explanation of the evidence is required to explain the interpretative process of vestiges, for example to indicate the type of layer/feature being reported. Finally, there will be evidential reflexive situations in which fieldworkers will require to include a short discussion of puzzling evidence when interpretations are ambiguous. Similarly, this analysis has allowed to identify the commonest forms of unreliable sources. For example, some sheet due the lack of an interpretation, an unjustified interpretation and an undiscussed reflexive testimony.

In general, this basic form of documental analysis belongs to a positivistic strategy to evaluate the credibility of testimonies. Thus, when a document results to be false or erroneous, it is discarded for being unreliable (Ginzburg 1991, 83). However, modern historiography has developed a more complex notion of documental analysis in which even false an unreliable testimonies can be a source of evidence when examined with a set of questions or "*we force them to speak, even against their own will*" (Bloch 1954, 53). This strategy has allowed to study the individual performance of fieldworkers and the longitudinal development of projects, for example. In consequence, documental analysis is no longer reduced to "*to purely and simply recording the words of our witnesses*" (Bloch 1954, 53). Besides, this necessarily opens a new question regarding the status of unreliable testimonies. Are these going to be discarded? However, considering the previous ideas, even unreliable sources can be useful if they are properly interrogated. For example, consider an inadequate description of a feature in which the fill sheet only mentions: "*fill of feature [744]. Pot rich fill*" (AH/GT-743), and the cut sheet describes an: "*oblong feature*" (AH/GT-744). Even such limited data can be reinterpreted as a rubbish pit considering the amount of pottery in the fill and the shape of the cut, and some similarities with other site features. Naturally, a more reliable reinterpretation of problematic testimonies would require confronting different textual and graphic testimonies. However, the main point is stablishing that even problematic records cannot be automatically considered as useless sources.

The second aspect is related to the issue of the apparent lack of reflexive testimonies in context sheets (Mickel 2015). Hence, MRG95's constitutes a strong proof of the presence of reflexive data in SCR. Moreover, this documental analysis allowed to differentiate between observational, persistent and evidential reflexive information, and the recurrent quality problems in evidential testimonies. However, two things should be noticed. The main core of evidential reflexivity are intentional testimonies from textual accounts and sketches, on the contrary the main core of observational reflexivity has been coded from unintentional testimonies as corrections in the stratigraphic sequence and sketches. The notion of unintentional testimony also emerged from the modern notion of documental analysis (Bloch 1954, 53). Then, when a document is analysed under this light, one can detect indirect evidence of the cultural manners of the Hebrews when reading the bible, instead of simply being read as a testimony of the deeds of Jesus (Ginzburg 2018, 158). Furthermore, there is a more important question that deserves further consideration because unlike reflexive narratives of discovery, corrections of sketches and diagrams in context sheets have a heuristic function during the definition of primary data. Hence, even if these testimonies are not fully useful in post-excavation, it is more difficult to define them as trivial from the point of view of the fieldworker. Besides, this evidence proves that even context sheets contain traces of the act of discovery.

Furthermore, this methodological revision of documental analysis allows to notice more clearly additional methodological problems with the principles of the reflexive method. Specifically, Hodder and his colleagues believe it's impossible to evaluate context sheets due its dry format and the lack of sufficient contextual information (Hodder 2005; Mickel 2015). In a large extent, this claim emulates some old scholar debates regarding the uselessness of dry medieval tales as historical sources, yet the evidential value of these testimonies changed when analytical strategies became more sophisticated (Ginzburg, 2012, p. 3). Besides, when reflexive archaeologist subscribe the impossibility to examine context sheets they contradict its own theoretical background based in the post-processual analogy between interpreting archaeological remains and documents (Hodder 2000b). Hence, the reflexive paradox is that of endorsing interpretative theory inspired in documental analysis but rejecting the possibility of examining context sheets, a form of interpretative exercise (Ginzburg 1999, 20). However, the most important problem with reflexive discourse is that they demand an additional sequence of diaries to create a background context that enables cross-examination between diaries and sheets. However, they fail to acknowledge that context sheets are already inserted in a complex contextual network formed by a sequence of sheets linked in multiple ways. Hence, they fail to acknowledge that cross-examination of context sheets is possible in the absence of diaries. In sum, the reflexive stance is attached to a very limited view of documental analysis in which they fail to acknowledge that even the most fragmentary testimonies can be valuable sources when properly interrogated. And instead, they adopt a rather naive position in which analysis is possible by the number of sources, not the strategies of interrogation. Evidently, this is not true because even the best sources will be silent in unqualified hands.

Finally, the documental analysis allowed to understand more clearly an important aspect of the recording strategy of SCR, namely the use of textual and visual testimonies in context sheets. This is an important difference with reflexive diaries which are primarily textual. Overall, there are two types of relation between textual and visual testimony in context sheets. First, there are numerous examples in which sketches and textual testimony work together to explain operative decision (Figure 4.8), technical mistakes (Figure 4.9) and interpretative decisions (Figure 4.10). Similarly, sketches can complement textual testimony to illustrate reflexive aspects like ambiguities in layers (Figure 4.11), give indirect testimony of episodes of the act of discovery (Figure 4.18 and Figure 4.21), and less commonly to illustrate reflexive discussion (Figure 4.25 and Figure 4.26). But more importantly is noticing that most of these testimonies contribute to contextualize the credibility of primary data.

Secondly, there are cases when textual and visual testimony work together as evidential support. For instance, when fieldworkers are very sure of the structural character of remains, sketches provide the main proof of the observed thing whereas texts basically describe its attributes (Figure 4.2). However, when diggers require to give account of an interpretative situation, textual and visual testimony describe relevant evidence, although the explanation is generally textually recorded (Example 4.4 and Example 4.8). Functional and formation processes are two areas where testimony is primary textual. However, there might be complementary use of textual and visual testimony if one considers the content of various related sheets (Figure 4.1, Figure 4.2 and Figure 4.3). As explained before, textual testimonies of spatial information are not very clear whereas visual testimonies are not very frequent (Example 4.12). Hence, this is an aspect that could benefit much more from sketching.

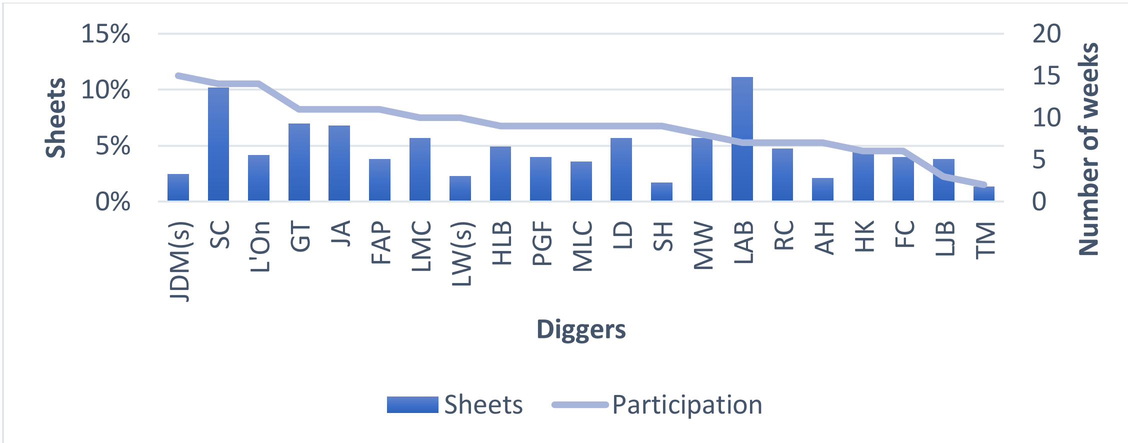
4.7 The performance of fieldworkers

The distribution of claims by fieldworker was examined to investigate some aspects of individual performance which could be explained by differences in skill and participation in the project. In general, all the participants at MRG95 were professional diggers, except for two supervisors (JDM and LW) that occasionally participated in recording labours. However, the most interesting difference was the diversity of ranges in project participation for professional diggers produced by frequent staff changes over a fifteen week period (Table 4.1). These ranges were reconstructed with an attendance lists available at MOLA's archive and checked with dates in context sheets. Table 4.1 doesn't list all the diggers that participated in the project, but only those included in the sample, but these represent most of the team. The exact number of participants can be checked in the published report (Seeley and Drummond-Murray 2005, xiv).

Digger	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15
JDM (s)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
SC		X	X	X	X	X	X	X	X	X	X	X	X	X	X
L'On		X	X	X	X	X	X	X	X	X	X	X	X	X	X
GT					X	X	X	X	X	X	X	X	X	X	X
JA					X	X	X	X	X	X	X	X	X	X	X
FAP					X	X	X	X	X	X	X	X	X	X	X
LMC			X	X	X	X	X	X	X	X	X	X			
LW (s)					X	X	X	X	X	X	X	X	X	X	
HLB					X	X	X	X	X	X	X	X	X		
PGF		X	X	X	X	X	X	X	X	X					
MLC		X	X	X	X	X	X	X	X	X					
LD							X	X	X	X	X	X	X	X	X
SH							X	X	X	X	X	X	X	X	X
MW					X	X	X	X	X	X	X	X			
LAB									X	X	X	X	X	X	X
RC	X	X	X	X	X	X	X								
AH				X	X	X	X	X	X	X					
HK			X	X	X	X	X	X							
FC	X	X	X	X	X	X									
LJB				X	X	X									
TM	X	X													

Table 4.1. Participation in the project by digger. MRG95. This table describes the participation of fieldworkers in relation to the number of weeks working in the site. Each row indicates in colour the week when a member joined and left the project. Fieldworkers that joined approximately at the same time are highlighted with a similar colour.

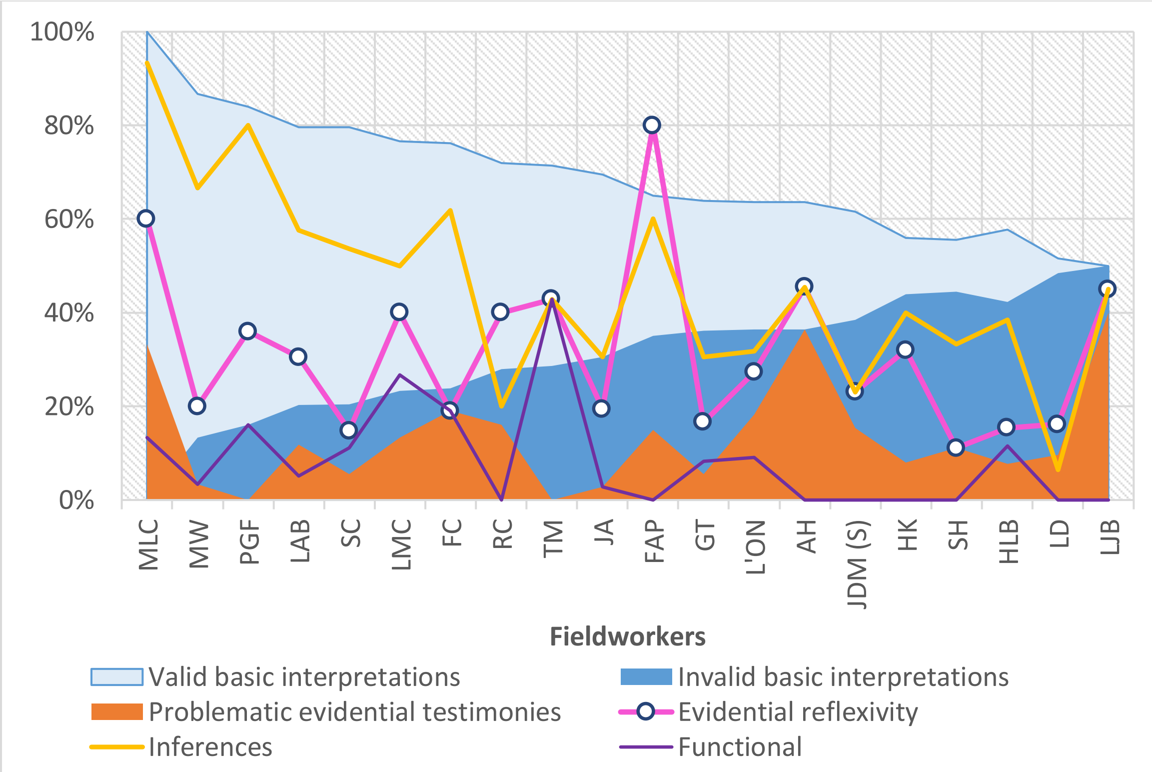
Additionally, Graphic 4.5 shows the distribution of the sample by fieldworker and extend of participation. Overall, this was one of the most difficult variables to control during sampling and the graphic shows that some diggers with shorter participation had a larger subsample (LAB) and others with longer participation a smaller one (LO'N). However, these are also related to the type and size of excavated deposits. Hence, to reduce representativeness problems, quantitative data was calculated independently for each participant.



Graphic 4.5. Distribution of the sample by digger. MRG95. This graphic illustrates the proportion of context sheets recorded by each participatn in relation to the extent of his participation in the project.

Graphic 4.6 shows the main results of the individual analysis and includes the distribution of valid and invalid testimonies by fieldworker. These results clearly show an inverse relation in the distribution of reliable and unreliable documents. In other words, the diggers that more often respect the standards of good practice are less prone to produce unreliable sources. Overall, this data seems to be the most indicative evidence of skill differences, independently of the extend of participation and the type of excavated archaeology. Furthermore, the values of valid data correlate with the proportion of inferences and functional claims. This evidence indicates that the more skilled fieldworkers also reach higher levels of interpretation more often.

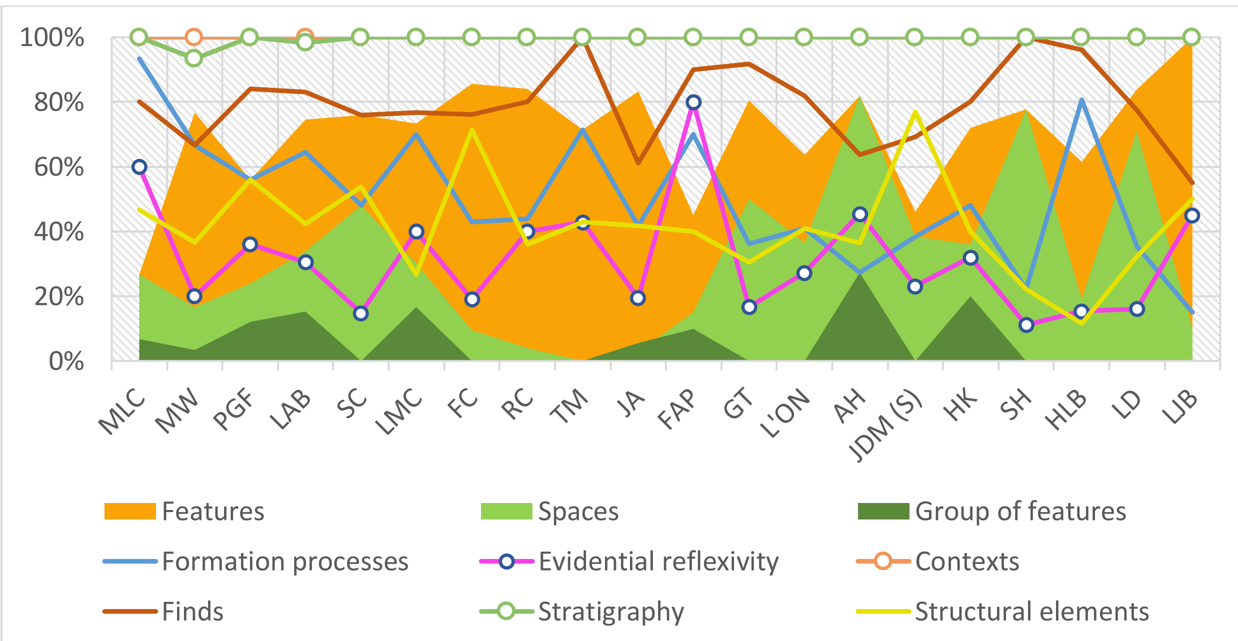
Graphic 4.6 also illustrates the values for evidential reflexivity which shows a relatively homogenous distribution, ranging between 20%-40% for every participant. In some extent, this distribution can be explained by having a trained workforce capable of developing a systematic reflexive behaviour. However, FAP stands out as one of the most reflexive fieldworkers because her records usually report ambiguities during the investigation of a big feature with fragmentary structural evidence (Example 4.15). On the contrary, SC and JA have lower reflexive values which are probably explained by a longer involvement with better preserved archaeology. Hence, such variations in reflexive data seem to be responding to contextual factors like the type of archaeology being excavated. However, the graphic also describes variations in the quality of evidential data which are better explained by individual skill.



Graphic 4.6. Digger's performance. MRG95a. This graphic illustrates the percentage of valid and invalid basic interpretations by fieldworker. Basic interpretations include context, structural and formation porcess data. Additionally, the graphic illustrates the distribution of problematic reflexive testimonies, relevant reflexive claims, inferences and functional claims.

Furthermore, Graphic 4.7 shows the distribution of additional types of claims with distinctive patterns. For example, context and stratigraphic data show a horizontal line in the top which represents the most systematically recorded types of data. Additionally, there are some irregular trends in the middle that seem to be responding to the type of archaeology excavated by each digger (finds, structural, formation process and feature information) which are very similar to the distribution of evidential reflexivity. Nevertheless, one can still observe certain homogeneity in the interpretative practice of fieldworkers as all of them cover the same areas. Altogether, the variability of information recorded by each participant represents the best evidence of a more 'democratic' interpretative practice based in a systematic behaviour independently of the extent of participation in the project. However, one must also take account that systematicity doesn't guarantee the quality and credibility of primary data, which are mainly regulated by skill.

There is only one area of where uniformity is interrupted, specifically in the distribution of spatial data. This trend shows very low values for five fieldworkers, including three diggers that worked in the earliest stage of the project {FC, RC and TM} which seem to have had less chances to accumulate enough background knowledge to have a wider understanding of the site. On the contrary, the low proportion of spatial claims for FAP and JA is explained by the type of archaeology that they investigated, specifically a huge ditch (Example 4.15). A similar explanation seems applicable for HLB who also worked in the big ditch and LJB that mostly worked in the stakeholes area (Figure 4.5). In this way, the distribution of spatial data represents good evidence of the incidence of contextual factors in the individual performance of fieldworkers.



Graphic 4.7. Digger's performance. MRG95b. This graphic illustrates the distribution of different types of claims by fielworker.

In general context sheets are clear testimonies of individual performance describing the observational and interpretative performance of fieldworkers (Figure 4.1). However, there are some atypical documents in which one can observe evidence of a collaborative record. One of the clearest examples is the initial description of layer [1010] by PGF, that later was updated by MLC (Figure 4.18). However, the soundest evidence of collaborative work can be observed when examining various related sheets. For example, a massive ditch [546] was initially defined by one digger (HLB-540, HLB-546) and subsequent sheets describe related elements to this ditch defined by other participants as in Example 4.15. However, the most interesting characteristic of this process is the way in which subsequent data are interpreted in relation to previous information. For instance, the interpretation of a posthole (FAP-808) is made in relation to the observations made by other participants (HLB-546 and JA-753). A similar example of collaborative work can be observed in Example 4.14.

Example 4.15

"Fill to ditch [546]. Fill was made up of many deposits with finely sorted alluvial deposits towards the bottom. At the base of [546]" (HLB-540).

"Ditch cut. Unclear as to ditch's purpose, may be defensive. Possible drainage into Walbrook tributary" (HLB-546)

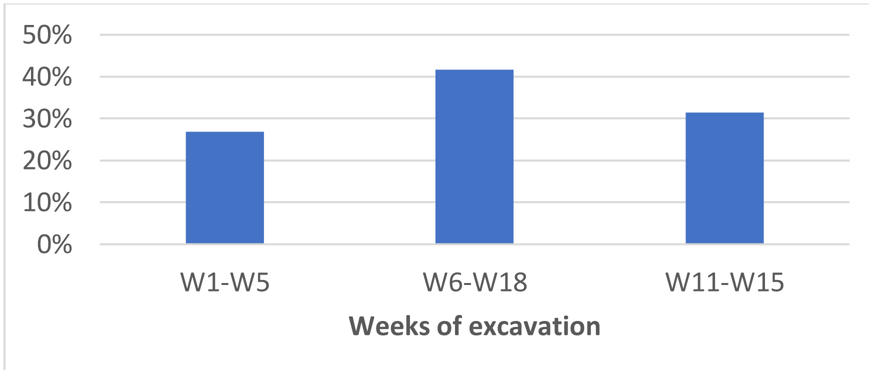
Cut of 6 stakeholes. They are all grouped on this context as they are associated with cut [546]. Are more than likely associated with ditch [546]. All except c) are located on west edge/side of ditch [546]" (JA-753)

Posthole/stakehole. Possibly associated with ditch [546], but more likely that this stake-hole is part of a different feature. Some of the characteristics of the stakehole are different to those associated with the ditch. This is regular in shape with a stake void larger in width and with a separate fill. This stake continues down through several layers and bottoms in layer [1105]" (FAP-808).

Unlike reflexive diaries, context sheets do not include testimonies of interactivity among fieldworkers. However, consistent stratigraphic corrections within testimonies elaborated by different agents are sound evidence that some form of communication was required. Independently of this, the most interesting feature of collaborative evidence is that testimonies an interpretive chain that connects observation and data produced by multiple agents. Thereby, even if diggers act as independent or subjective agents of knowledge, they do not produce subjective views of the site. In that sense, it could be said that diggers may excavate vestiges, but teams interpret sites.

4.8 Longitudinal analysis

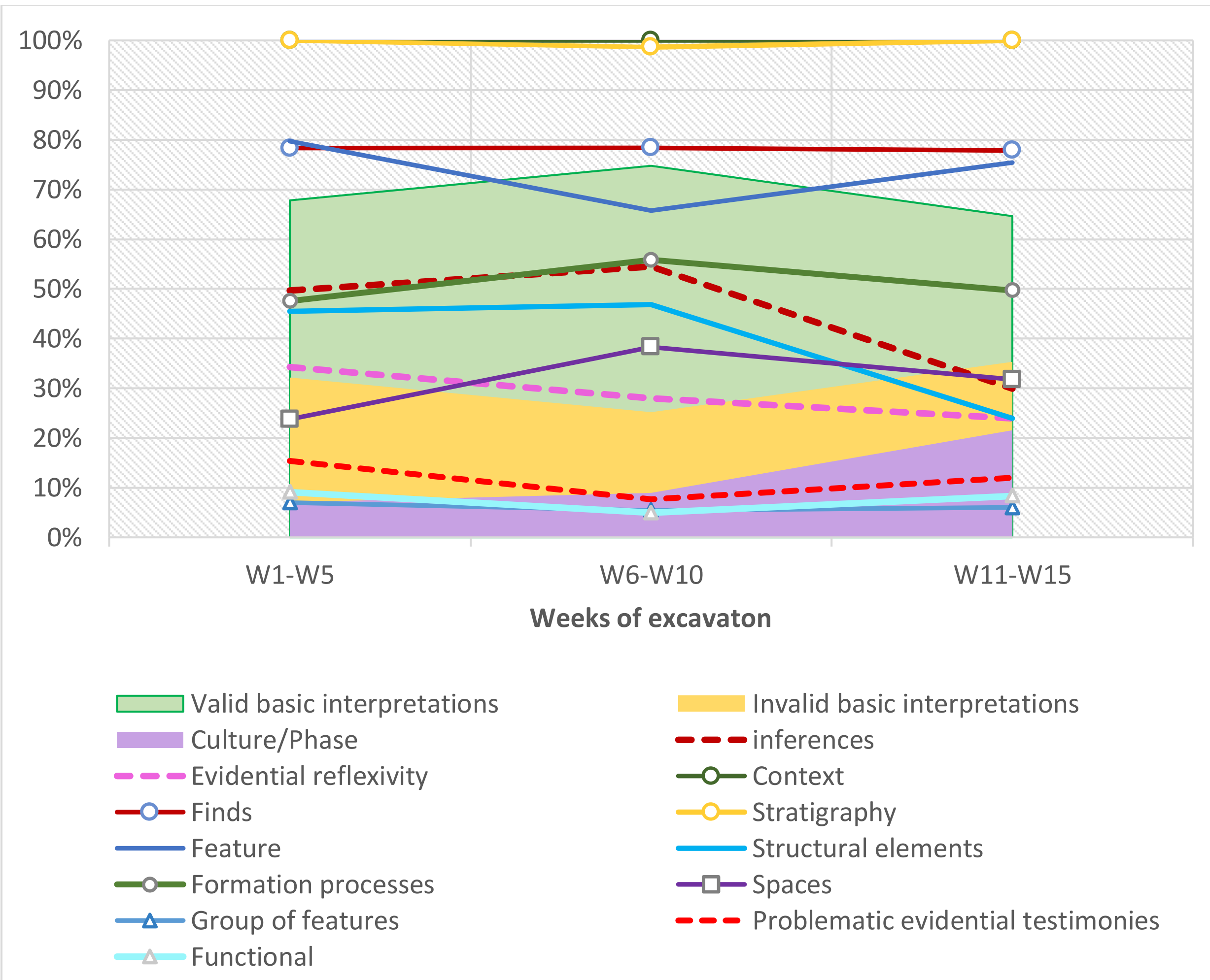
The excavation of MRG95 lasted for 15 weeks between the 15 of February and the 28 of May 1999, having a similar duration to B49's dig. For that reason, the sample was divided into three arbitrary stages for its longitudinal analysis. Each one of these phases covers five weeks and approximately represent a month’s work, this decision produced an uneven division of the sample and for this reason, the distribution of claims was calculated independently for each phase.



Graphic 4.8. Distribution of the sample into three arbitrary stages. MRG95.

Graphic 4.9 illustrates the longitudinal distribution of claims for the most recurrent information. In general, the purpose of the longitudinal analysis was observing differences and similarities in the interpretative practice of the project across time. Graphic 4.9 shows that the most frequent types of data (context, stratigraphy, finds and formation process) are continuously recorded in every phase of the project. Likewise, functional and groups of features are continuously too, although less frequently. Structural and feature claims have a less horizontal pattern, however when structural data decrease, feature information augments, therefore representing a regular investigation of these aspects. Overall, these results confirm that there are many aspects of excavation work that are regularly executed, independently of the phase of the project and in some extent the skill of the fieldworkers.

In contrast, spatial and culture/phase information shows a continuous increment. This evidence suggests that capturing such primary data is a function of time which most likely contributes for the accumulation of background knowledge that allows making more complex inferences. Even though the quality of spatial and culture/phase data is not very good along the whole project. On the contrary, the number of valid testimonies and inferences shows a regular start followed by an an important decrement in the last weeks which is correlated with an increment of invalid testimonies. Similarly, there is a slight but continuous decrement in the production of reflexive data although the production of problematic reflexive testimonies is very regular. This evidence suggests that there might be positive and negative effects in interpretative practice with the course of time. Among the positive aspects is the accumulation of background information for the interpretation of more complex data. Whilst the most negative aspect is a decrement in the reliability of data which most likely is correlated with the pressure of finishing a contract site on schedule. Thereby, although the analysis of digger's performance indicates that the quality of records largely depends on individual skill, the longitudinal analysis suggests that even the most skilled practitioners can be less efficient under pressure.



Graphic 4.9. Longitudinal distribution. MRG95. This graphic shows the distribution of valid and invalid interpretations, and various types of claims over the three phases of the project. Each phase is defined by a period of five weeks.

4.9 General observations

The internal analysis of MRG95 has provided a general view of the recording strategy and the interpretative process of the project by studying the variability, the credibility and the relevance of data. Additionally, the analysis has shed some light regarding the incidence of individual skill, background knowledge and the character of remains in the investigation of the site.

Recording strategy: Very often theoretical discourse characterizes SCR as rigid system mostly due the standardized format of context sheets. In principle, this argument can be partially rejected by considering the variability of formats (layer/cut, masonry, timber and burial) but ultimately the structure of context sheets is very similar. However, the soundest evidence of the flexibility of the system come from the various examples in which fieldworkers adapt recording rules to their interpretive needs and include crucial data that is not required by a predetermined field. Another crucial element of the recording strategy is its sequential nature that enables to link events. Thus, despite the interpretative process is fragmented into smaller pieces of information the contextual relations of data are never lost. Besides, the sequential production of records also facilitates to capture the conditional aspects of excavation like the fluidity of interpretations. Finally, another positive aspect of the recording strategy is the use of textual and visual testimonies which work together to contextualize the credibility of data and provide evidential support of interpretative data. For these reasons, it is difficult to agree that SCR represent a rigid recording system. There is only one area in which the format negatively affects the quality of spatial data because diggers generally tick a predetermined option whereas explanatory accounts and sketches are not very common.

The interpretive process: Another recurrent critic against SCR is an alleged restrained variability of data in context sheets. In general, this idea should not get much credit because context sheets are designed to capture multiples types descriptive, stratigraphic, interpretative and spatial information. Moreover, the analysis of context sheets also proved the capture of relevant non-interpretative and reflexive data. Furthermore, the internal demonstrated a wide diversity of interpretive and reflexive situations produced by the variability of types in layers, features, depositional processes and levels of preservation of remains. The qualitative analysis of primary testimonies also illustrated the coexistence of various standards for reporting different epistemic situations. For example, the report of a well-preserved structure only requires a description whereas giving account of fragmentary remains and sedimentary layers generally requires an explanatory account of the interpretive process, and a reflexive testimony usually requires a brief argumentative discussion. The analysis also confirmed the most recurrent quality problems in context sheets are the lack of interpretation, an unjustified interpretation and an undiscussed reflexive statement, but none of these problems are attributable to the recording method. Instead, they are product of human error. Finally, the internal analysis demonstrated the existence of an important epistemic change between the evaluation and excavation phases of the project, because the last contributed to a more detailed mapping of the industrial activities. Similarly, there is another epistemic change in post-excavation work that contributed to the interpretation of a building and the constructive phases of the site. However, it’s also clear the excavation team was less successful to cover these aspects, especially the interpretation of a building, due the fragmentary character of evidence. In sum, MRG95 presents a clear example of a rich interpretative practice in commercial field practice.

Reflexivity. The analysis of reflexive data in context sheets confirmed the distinction among observational, persistent and evidential reflexivity. Although, the focus of SCR is evidential reflexivity. Besides, it’s clear that a lot of evidential reflexivity at MRG95 was triggered by fragmentary remains of structural evidence and marginal evidence of sedimentary layers. Another characteristic of reflexive recording at MRG95 is the use of textual and graphic elements to describe reflexive situations. Either to capture episodes of discovery or to describe and discuss cases of puzzling. A recurrent problem with evidential testimonies is the lack of an argumentative discussion which indicates that even reflexive situations must be regulated by standards of good practice.

Non-Interpretative claims. The analysis of circumstantial and decision claims showed the existence of two classes of non-interpretative data. First, a systematic information in context sheets describing the general observational conditions and the technical procedures implemented onsite. Additionally, there is a small number of sheets that include additional information for specific situations. For instance, if a layer was partially excavated due safety protocol or a technical error during the definition of a layer. Altogether, these testimonies provide a contextual background of specific observational and operative actions which also contributes to contextualize the credibility of data. Finally, it was observed that sampling data is limited, but interpretative information generally provided an important contextual background to understand the rationale of sampling activities.

Individual and longitudinal analysis.

The analysis of the performance of fieldworkers showed that almost every participant of the project recorded the same variability of data, this seems to be an important proof of the benefits of a systematic approach by a trained workforce. Although, the credibility of testimonies also demonstrated important differences in the skill of participants. The longitudinal analysis confirmed this systematic approach of interpretation, except for spatial data which is more sensible to the accumulation of background knowledge in time. This anomaly also was indicated by the individual analysis which provided some evidence that diggers with shorter participation had less chances to make spatial interpretation. Finally, another important contribution of the longitudinal analysis was providing evidence of a decrement in the quality of records in the last weeks of the project, which suggest that time pressure might be another important factor affecting the quality of records.

Chapter 5. The feature system

Archaeology does not exist in a vacuum; every visitor, every supporter as well every volunteer and professional excavator leaves a mark on the record for better or worse. Martin Carver, 2016, 42.

This chapter presents the results of the third case study which examines the use of the feature system in the Tarbat Discovery Programme. The chapter begins with a brief account of the methodological principles of the feature system, this section also includes a description of the formats for the context and feature sheets used in this project (5.1). Afterwards, there is a description of the project (5.2) and the site (5.3) which altogether describe the epistemic context of the investigation. In broad terms, the Tarbat Discovery Programme can be described as a long-term research project that excavated a landscape site containing the remains of an early medieval monastery. The site was excavated by a mixed team of professional archaeologists and trainees. The chapter also includes a brief description of the sampling process of the sequences of context and feature sheets (5.4). The records of the Tarbat Discovery Programme are under the custody of Historic Environment of Scotland (HES) in Edinburgh, Scotland.

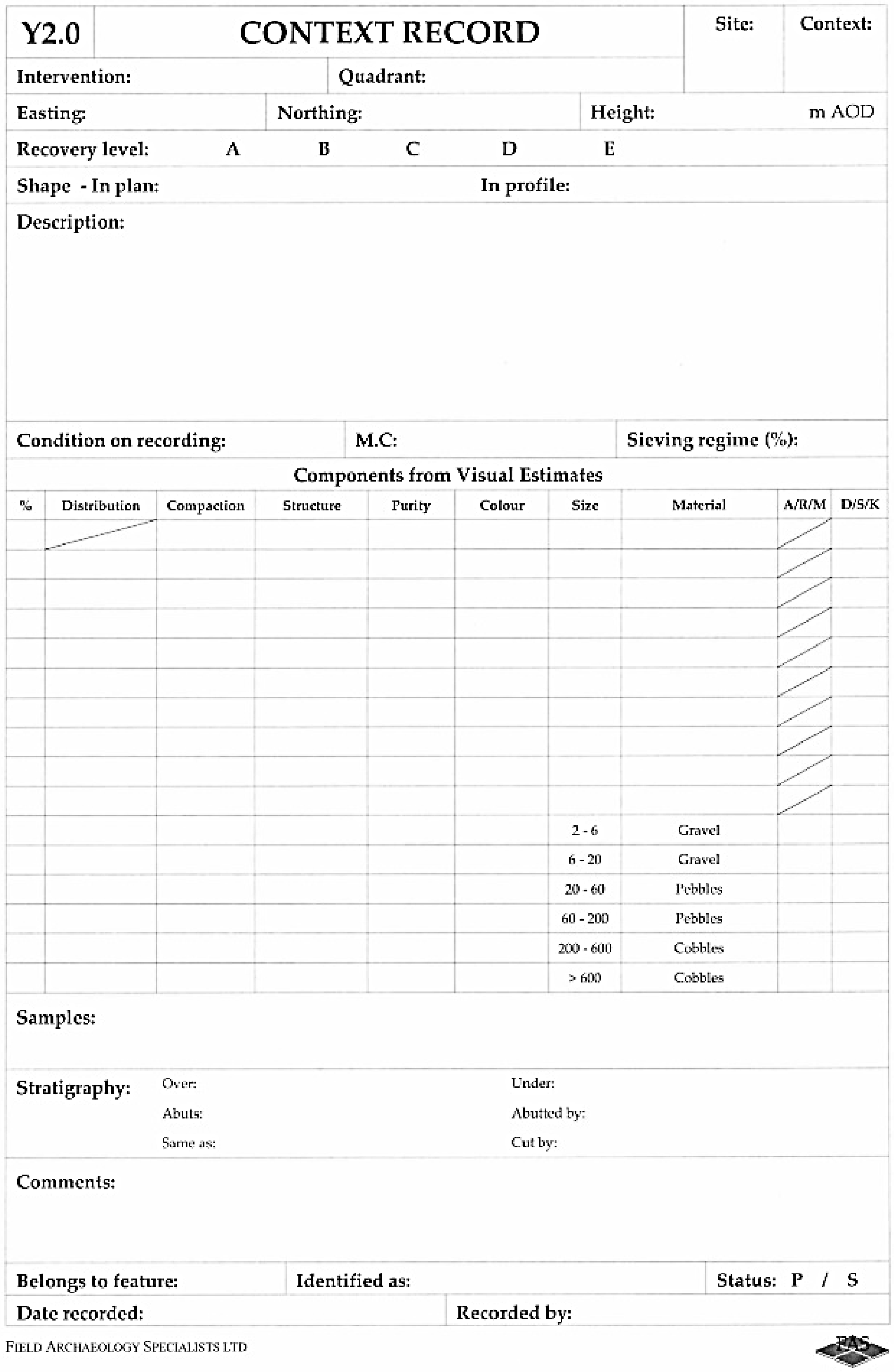
As explaine before, Martin Carver has become one of the keenest promoters of the feature system in recent years. However, there are some inconistencies in his thought because in one hand, he believes that introducing an aditional sequence of feature sheets automatically will improve the quality of primary data but at the same time he acknowldges that the perfrmance of fieldworkers is a determinant factor in the quality of data. Hence, the internal analysis of this case study explores the variability, credibility and relenacy of information (5.5). Altogether this information will be useful to identify the strengths and problems in the design and the implementation of the feature system. The analysis is complemented with an examination of the distribution of claims by participant (5.6) and a longitudinal distribution of claims (5.7) which will contribute with important evidence to demonstrates the incidence of contextual factors in field investigation. The chapter finishes with a recapitulation of the main observation derived from the analysis of the final case study (5.8).

5.1 The recording system

The feature system seems to be derived from an alternative tradition of fieldwork related to the explorations of landscape archaeology (Hammer 2000). For example, the methodological developments at Yeavering exploration by Brian Hope-Taylor seem to be an important antecedent to the development of the feature system (Carver 2016, 31). But in general, the history of the feature system has not received the same attention than the methodological developments for urban explorations. In recent years, Martin Carver has been a keen supporter of this alternative method (Carver 2009), specifically he has implemented the feature system in two research project in early medieval sites. The first was an exploration in the famous Anglo-Saxon cemetery of Sutton Hoo (Carver 2005) and the most recent was an exploration in a monastic settlement in Portmahomack, Scotland (Carver,Garner-Lahire and Spall 2016).

The feature system is based in the idea that a wider conceptual toolbox will facilitate the interpretation of a wider variability of entities like *contexts*, *features* and *groups of features* (Carver 2009, 121). For this reason, the feature system includes various sequences of formats hierarchically organized for the description of deferent types of information. In principle, one sequence for describing context and another for describing features. However, this procedure can be repeated by introducing an additional sequence for describing groups of features. In this way, the feature system also links a group of events of a sequence to an event from another, for instance a group of context sheets will be related to a feature sheet. However, one of the goals of the feature sheet is to provide a more synthetic testimony that facilitates analysis, unlike SCR where feature information is dispersed across many related documents and therefore analysis might be more complicated. Nevertheless, it should be kept in mind that contexts will not be related to a feature every time, and these events will be called floating contexts (FC).

The examined formats of the feature and the context sheet used in the Tarbat Discovery Programme were designed by a small company Field Archaeological Services (FAS) of which Carver himself was a founder. In broad terms, Tarbat's formats are standardized sheets which show similarities and differences with MOLA's formats. For example, Tarbat's system includes diverse formats of context sheet for the description of deposits (Y2.0), timber (Y2.1), skeletons (Y2.2), coffins (Y2.3) and masonry layers (Y2.4). Even-though the structure of all of them is very similar. However, an important characteristic of Tarbat's method is the absence of a cut sheet, being this a crucial difference with MOLA's system.



Y2.0	CONTEXT RECORD	Site:	Context:

Intervention: | Quadrant:

Easting: | Northing: | Height: m AOD

Recovery level: A B C D E

Shape - In plan: | In profile:

Description:

Condition on recording: | M.C: | Sieving regime (%):

Components from Visual Estimates

%	Distribution	Compaction	Structure	Purity	Colour	Size	Material	A/R/M	D/S/K
						2 - 6	Gravel		
						6 - 20	Gravel		
						20 - 60	Pebbles		
						60 - 200	Pebbles		
						200 - 600	Cobbles		
						> 600	Cobbles		

Samples:

Stratigraphy: Over: | Under:
Abuts: | Abutted by:
Same as: | Cut by:

Comments:

Belongs to feature: | Identified as: | Status: P / S

Date recorded: | Recorded by:

FIELD ARCHAEOLOGY SPECIALISTS LTD

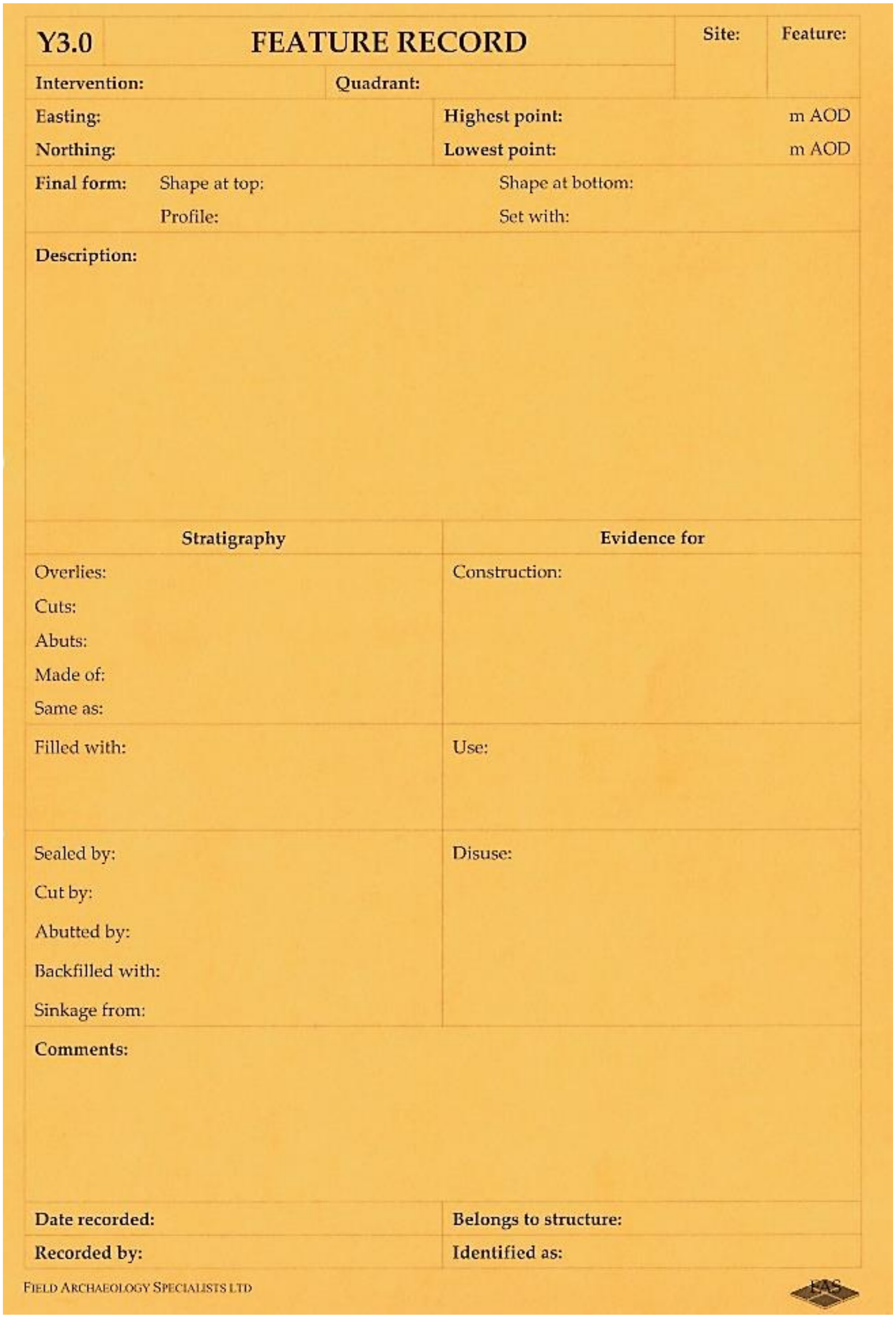
FAS

Figure 5.1. Context sheet of the recording system used at Tarbat. Courtesy of FAS.

The context sheet (Y2.0) is organized in different sections which largely emulates the structure of any context sheet which includes two main areas. One for descriptive data and another for interpretative information (Figure 5.1). Tarbat's context sheet includes two big fields for layer description: *Description* and *Visual Estimates*, which occupy an important portion of the sheet. Additionally, there are some complementary field for describing layer's *Shape in plan* and *in profile*, and colour (*MC*). The sheet also includes some boxes to annotate information about observational circumstances (*Condition on recording)* and operative decisions like *Sieving regime*, *Sample* and *Recovery levels* which include five predetermined options (A-F). On the low area, there is a section for *Stratigraphy*, which in fact registers physical contacts (over, abuts, under and so forth) and below there is a space for *Comments* which commonly contained interpretative data. Almost down the bottom there are two additional fields, one to indicate whether the context *Belongs to a feature* where diggers must record a feature numbe. The other is a box to introduce the type of layer being described (*Identified as*). Finally, there is a small box tagged as *Status* which includes two predetermined options, one for primary deposits (P) and another for secondary layers (S). The format also includes multiple boxes for location references (*Sit*e, *Intervention*, *Quadrant* and *Coordinates*) and identification of data (*Context*, *Date* and *Recorder*) in the top and bottom rows of the sheet, but none of this was analyses. In this way, Tarbat's context sheet has important differences with MOLA's format. For instance, the lack of a cut's sheet, the lack of a Harris matrix and the lack of a find's section. Another characteristic of Tarbat's context sheet is being one sided and the lack of a drawing section, nevertheless, some sheets included additional comments and sketches in the blank side of sheets. The implications of these characteristics will be explained in the course of this chapter.

The <u>feature sheet (Y3.0)</u> shows a similar structure and it includes sections for location and identification data on the top and bottom, and areas for feature *description* and *interpretative comments* (Figure 5.2). The feature sheet also includes a section for introducing a brief description of the type feature (*Identified as*). The most distinctive element of the feature sheet are the two columns located in the centre. One for additional *Stratigraphy* data and another for *Evidence for* subdivided into *Construction*, *Use* and *Disuse* rows. These rows are designed to describe the formation processes and functional aspects of features, which are correlated with stratigraphic sections, thereby the upper row for stratigraphy is designed to record physical relations produced by constructive events (*Overlies*, *Cuts*, *Abuts*, *Made of* and *Same as*), the middle row '*Filled with*' aligns with *Use* and finally the lower section (*Sealed by*, *Cut by*, *Backfilled with*) do the same with *Disuse*. Nevertheless, there are some conceptual problems with these entries, because some of these indicate group relations (*Made of*, *Filled with* and *Backfilled with)* rather than stratigraphic data. Still, there is a chronological arrangement of this data if one considers that that constructions events (*made of*) are earlier than use events (*filled with*) and disuse events (*backfilled with*). Another confusing aspect is the presence of cutting relations but no actual space for cuts description, yet as it shall be explained, cuts information is commonly recorded within feature descriptions.

Additionally, the feature format contains two small fields to indicate relations between groups of features. One is in the top of the sheet labelled '*Set with'*, the other is in the bottom tagged as '*Belongs to structure'*. The last seems to be specially designed to indicate when a group of features defines a building structure. FAS system includes an additional format, the Structure sheet (Y4) to describe groups of features, nevertheless, any record of this type of format was found in Tarbat's archives, then it will not be considered here. A remarkable difference between the context and the feature sheet is their colour and although no further consideration of this will be made, this attribute was quite useful during the process of analysis and presumably onsite. Finally, it should be said that Carver also supports the use of excavation diaries, however unlike the reflexive method, diary records are restrained to site supervisors (Carver 2009, 144). Tarbat's archive includes site diaries for every excavation sector, however diary recording was not very systematic. For example, the diaries of Sector 2 only cover three seasons out of eleven, therefore being unsuitable for documental analysis, nevertheless some minor observation about these documents will be made afterwards.



Y3.0 **FEATURE RECORD**

Site: | Feature:

Intervention: | Quadrant:

Easting: | Highest point: m AOD

Northing: | Lowest point: m AOD

Final form: Shape at top: | Shape at bottom:

Profile: | Set with:

Description:

Stratigraphy	Evidence for
Overlies: Cuts: Abuts: Made of: Same as:	Construction:
Filled with:	Use:
Sealed by: Cut by: Abutted by: Backfilled with: Sinkage from:	Disuse:

Comments:

Date recorded:	Belongs to structure:
Recorded by:	Identified as:

FIELD ARCHAEOLOGY SPECIALISTS LTD

Figure 5.2. Feature sheet of the recording system used at Tarbat. Courtesy of FAS.

5.2 The project: The Tarbat Discovery Programme

The Tarbat Discovery Programme (TDP) was a research project for the investigation of the Pictish site of Portmahomack, located in the Tarbat peninsula of Scotland. The project emerged by a local initiative to investigate and preserve The Church of St. Colman (Carver 2016, xiv), although the investigations were directed by Martin Carver a professor of the University of York specialized in medieval archaeology. Fieldwork began in 1994 and finished in 2007, which included evaluation and excavation work in three large sectors: Sector 1 and Sector 2 explored areas of domestic and industrial activity, whereas Sector 4 investigated the remains of the Church of St. Colman (Figure 5.3).

The execution of the project was supported by the professional services of a small commercial unit (FAS-specialists). In addition, most seasons included the participation of trainees from the University of York and local volunteers (Carver 2016, xiv). In this way, the TDP was a long-term research program assisted by a commercial unit in a non-development situation. Each field season was followed by the publication of a preliminary report that summarized the main findings and preliminary interpretations of the site (Spall,Garner-Lahire and Carver 2012). But the main post-excavation program lasted between 2007-2016 which concluded with the publication of the final report (Carver,Garner-Lahire and Spall 2016). Overall, the TDP lasted for more than twenty years funded at various stages by different sponsors (Carver 2016, xv).

Another important characteristic of the project is its epistemic context. In general terms, the TDP has been described as the first exploration of a Pictish settlement, although pictish monuments had been investigated before (Carver 2016, 1). More specifically, the monastic character of the site was largely evident by the presence of the church and evaluation work provided some clues regarding the characteristics of deposits and the extension of the site. Hence, the purpose of the different excavation sectors was to investigate the range of social activities related to the monastery (Carver 2016, 27). This aim to match the research agenda with the background knowledge of the site forms one of the cornerstones of a wider methodological strategy that Carver defines as Evaluative archaeology (Carver 2009, 32).

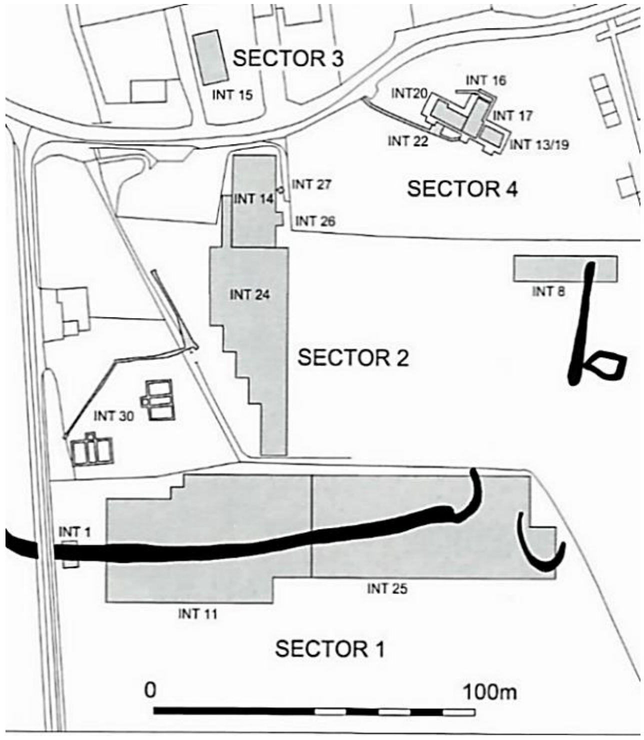


Figure 5.3. Excavation sectors at Tarbat (Copyright Tarbat Discovery Programme).

Figure 5.4. General view of Sector 2 (TRS2) (Copyright Tarbat Discovery Programme).

5.3 The site

Tarbat is an early medieval site with the characteristics of a landscape deposit as the archaeological remains extend more horizontally than vertically, and this factor was crucial in the strategy of excavation, especially in the size of excavation areas (Figure 5.4). Specifically, the archaeology of the site included a building church in Sector 4, which contained many human burials. On the contrary, Sector 1 and 2 were large areas of industrial activity with a wide diversity of positive and negative features. In broad terms, the survival of remains was relatively good, especially for stone structures.

Sector 2 was selected for the analysis of records due its wider variability of features across different occupation phases. However, the main occupation dates from the monastic phase which includes a large industrial area divided by a road. To east section included the traces of a timber building for '*vellum production'*(S9) and a yard workshop, whilst the east area contained the remains of a metalworking area and water management features. Likewise, the north and south areas of Sector 2 included additional features for water management (Figure 5.5). After the destruction the monastery, the site changed into a small village (Carver 2016, 64–75). The excavation of Sector 2 was part of the annual program between 1996-2007, except for 2002. This huge sector was divided in two smaller areas. Intervention 14 (INT14) which covered the norther quadrants and Intervention 24 (INT24) for the southern, but the extension of the full grid covered 100m long and 30m wide, with depth variations between 30 cm and 2m at its deepest (Carver 2016, 64). Thereby, being the largest and the longest exploration among the examined projects.

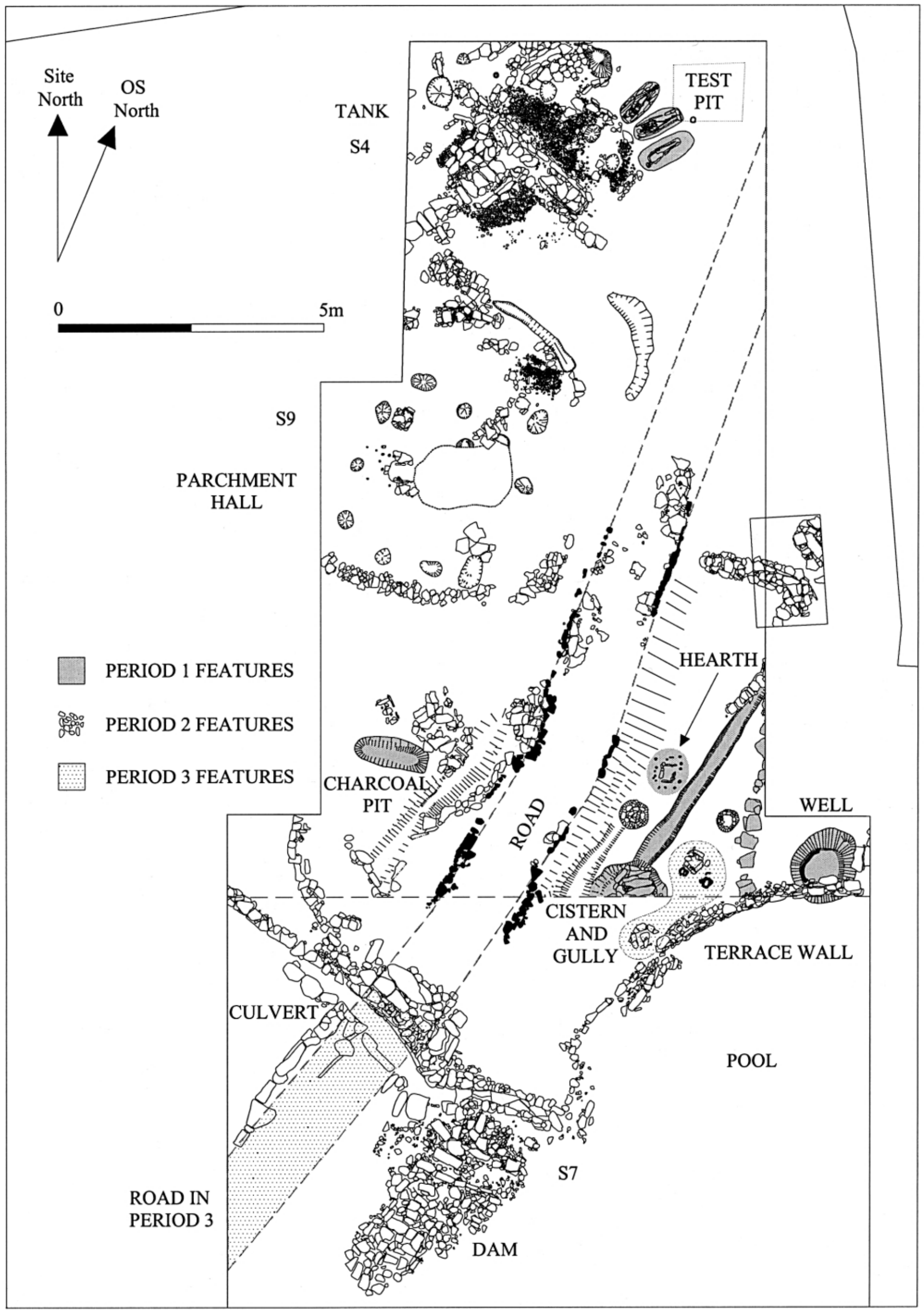
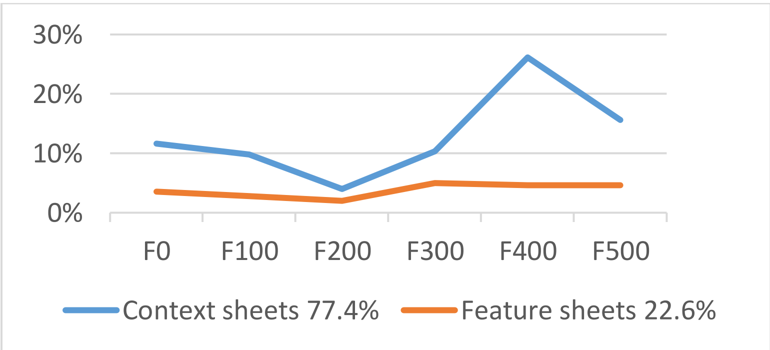


Figure 5.5. General plan of the monastic occupation in Sector 2 (Copyright Tarbat Discovery Programme).

5.4 The sample

Tarbat's sample was collected following the same criteria used for London's sequence, except for minor modification. Specifically, Tarbat sample selected sets of grouped sheets, that is context sheets (CS) associated with a feature sheet (FS), without considering floating contexts (FC). This selection was designed to investigate more closely the interpretation of related layers and features and compare the context in both types of sheets. The sequence of context sheets from Tarbat runs from C1000 to C3655, whereas the feature sequence runs between F1-F581. The sample includes 501 sheets: 388 context sheets (CS) and 113 feature sheets (FS) which represent 20% of the features from Sector 2. As in former cases, the sample was designed to cover various participants and stages of the exploration. Thereby, the number of features was evenly distributed across the sequence, but the number of context sheets was more difficult to control (Graphic 5.1). The sample included two or three examples of each type of feature (postholes, hearths, middens, ditches, walls and so forth) aiming to identify the strengths and weaknesses when investigating similar and different types of features. Most of the sample was selected from the central area from Sector 2 because this area has the major concentration and variability of features (Figure 5.5). The sample covers features from all the periods of the site, yet there is a good number of the monastic period, which represents the most intensive occupation of the site. Tarbat's archive are under the custudy of Historic Environment of Scotland (HES) at the John Sinclair House in Edinburgh, Scotland. Sector 2 records are stored in the *General Collection, Excavations at Tarbat West Church* (551 359/76), specifically in boxes of *Intervention 14/24, Acc No: 2017/14,* which cover *Unit ID: 33109-33316*.

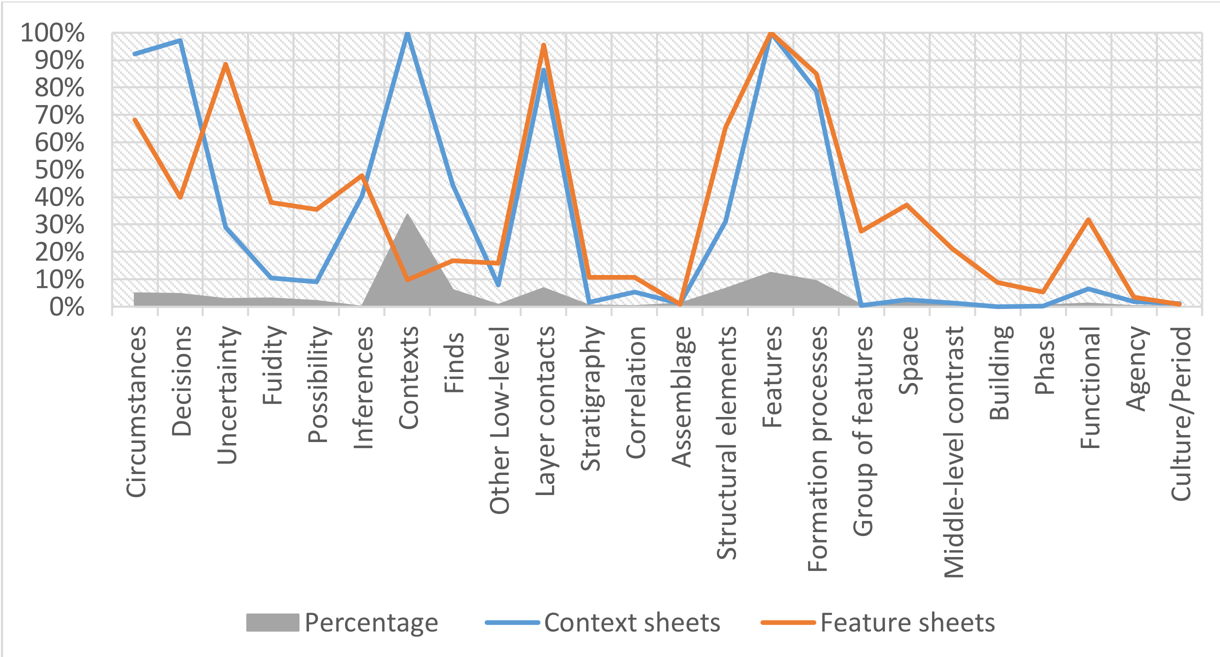


Graphic 5.1. Distribution of the sample. Tarbat.

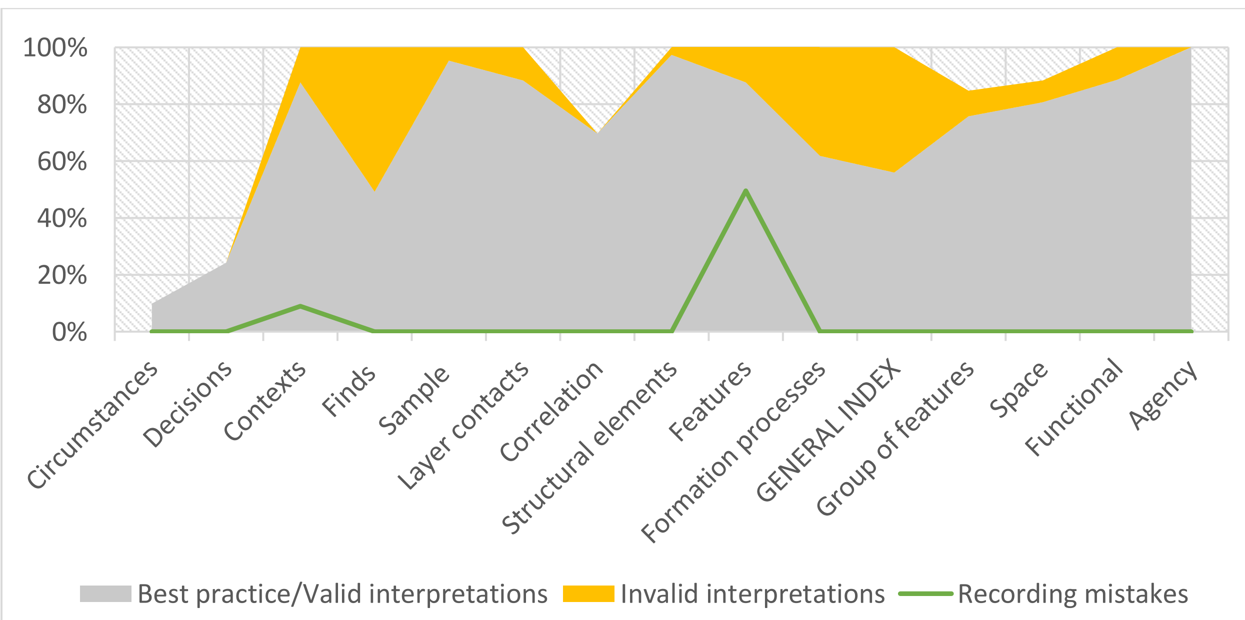
5.5 Results

The analysis of Tarbat records also had minor modifications in the quantification of data due the double sequence of context and feature sheets. Graphic 5.2 shows the frequency of claim in context and feature sheets, whereas percentage illustrates the distribution of information independently of the type of sheet. In this case, frequency will be the basis for internal analysis too, meanwhile percentual information will be used for the comparation among case studies. In general, Graphic 5.2 shows a wide variability of information with a similar distribution to previous case studies. For example, there is a higher frequency of layer, feature and structural information whereas spatial, building, phase and high-level claims less frequent. However, the presence of building and phase claims deserves a special mention. In this case, the high frequency of feature claims is explained by the sampling strategy, but more interesting is the asymmetric distribution of layer description in context sheets and the larger concentration of spatial and groups of feature claims in feature sheet. Whilst layer contacts, structural, feature and formation process information is similarly recorded in both formats. Finally, one can also observe a higher frequency of reflexive claims in feature sheets and a similar proportion of inferences in both type of sheets.

Graphic 5.3 shows the results of qualitative analysis of the credibility of interpretative data. As explained before, valid claims represent the proportion of reliable testimonies whereas invalid claims represent the proportion of unreliable data. These proportions have been calculated in relation to the frequency of each type claims independently of its location in context or feature sheets. This graphic also includes a general index that summarizes the total proportion of valid and invalid testimonies in basic interpretative data: context descriptions, structural data and formation processes. Finally, this chart includes an additional value that has been defined as recording mistakes, these represent minor recording failures without compromising the reliability of data.

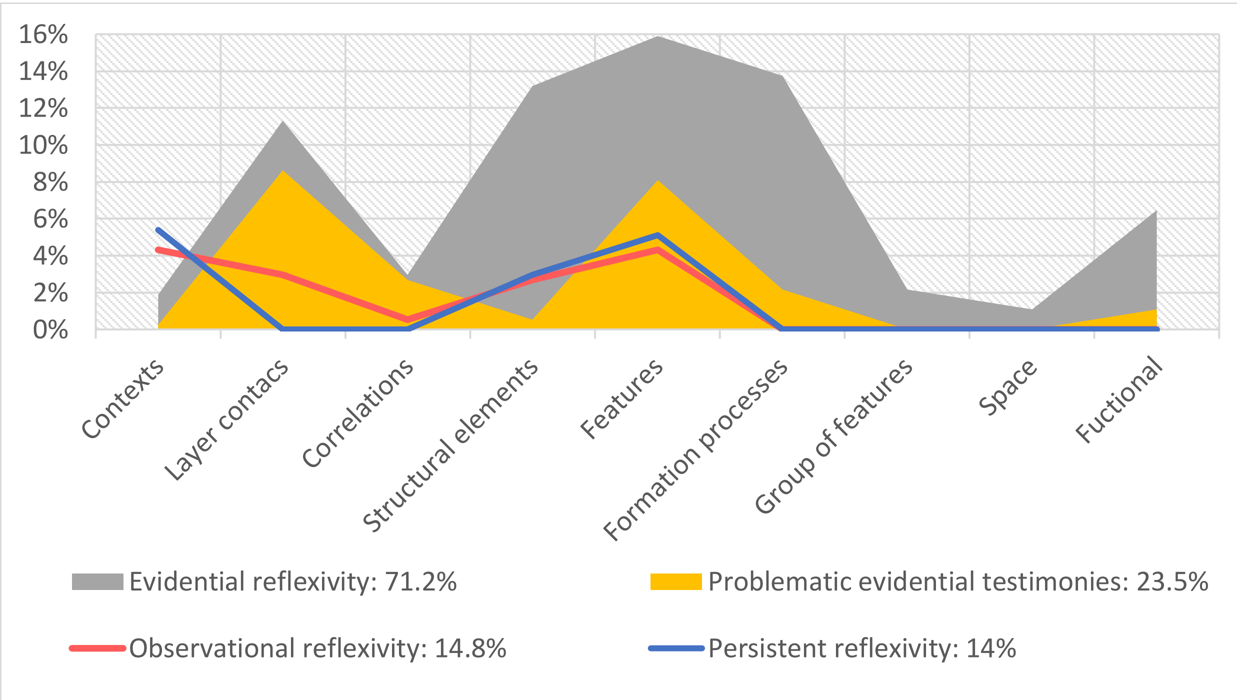


Graphic 5.2. General distribution by type. Tarbat. The graphic illustrates the frequency of claims in context and feature sheets. And the percentual distribution of data in the sample.



Graphic 5.3. Credibility of interpretive claims. Tarbat. The graphic illustrates the distribution of valid and invalid interpretative data for the total sample (General Index) and by type. Additionally, it includes the describes the proportion of best practice claims for circumstantial and decision claims. Finally, it depicts the proportion recording mistakes for some types of claims.

Finally, Graphic 5.4 describes the results after the classification of reflexive information into three categories: observational, evidential and persistent claims. This graphic shows an important emphasis of reflexive information in evidential data, even though there is a large amount of problematic evidential testimonies equate. The proportion of observational reflexivity is much lower but includes coded examples from intentional and unintentional testimonies.



Graphic 5.4. Reflexivity. Tarbat. The graphic illustrates the distribution of reflexive claims according to three qualitative categories (observational, evidential and persistent) and in relation to interpretive claims. Also, it describes the distribution of recording problems in evidential claims.

5.5.1 Non-Interpretative information

Information about observational circumstances and decisions was found in both context and feature sheets alike (Graphic 5.2). Although, the system is designed for the capture of this data in context sheet (Figure 5.1). Specifically, context sheets include two systematic pieces of non-interpretative information. First, *conditions of recording* generally describe the observational condition when removing a layer either as damp/wet or dry. Secondly, contexts sheets include different areas for describing operative aspect like recovery levels, which are selected accordingly if the layer had been dug with trowel (C) or mattock (B). In general, this information didn't seem very relevant except as a general account of the observational conditions and the technical procedures in the site. Contexts sheets also include standardized fields for sampling and collecting decisions, but these aspects will be discussed later.

Additionally, some context and feature sheets include further non-interpretative data the observational circumstances and operative procedures during the exploration remains. For example, many sheets describe a long gap between the definition and excavation of features: "*F18 was identified in TR96 but not fully defined and freed up for excavation until the beginning of TR99, when it was excavated at the beginning of the season*" (F18-CAS)[8]. This information is an indirect testimony of the slower pace of excavation at Tarbat which in many occasions could be influenced by the size of features (Figure 5.5). Nevertheless, another factor of the pace of exploration was the excavation strategy based in horizons, as indicated in some testimonies: "*F357 was defined at Horizon 3B after the removal of C1793*" (F359-LZ). Digging in horizons implies the definition of extensive surfaces that cover many types of archaeological remains which seem to be contemporary (Carver 2009, 117, 2016, 31).

Another characteristic of the project was the excavation of negative features, which commonly were defined in plan but explored in section. Hence, site records indicate when an account was based in the partial observation of something (

Figure 5.6). Tarbat sheets also include brief notes to describe when data collection was restrained by the limits of the excavation area: "*The original depth seen in section was a maximum of 0.17m. F195 also appears to continue across module D3 under the eastern section of INT24*" (F195-DW) or when a feature had been explored with atypical procedures for a specific reason: "*F166 was excavated out of sequence at its western limit to recover sculpture 1798*…" (F166-ST). Likewise, some sheet report operative errors in recording activities with no repercussions in the credibility of claims: "*because of a misunderstanding during the allocation of feature numbers, F383 appears as F382 on the section photographs*" (F383-LG).

[8] For referencing testimonies from primary records two are used. Testimonies from feature sheets are references with a feature number and the initials of the author (F18-CAS). Testimonies from context sheets are referenced with a context number and the initial of the author (C1490-CKH).

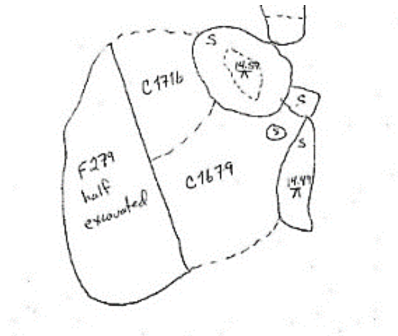


Figure 5.6. Sketch of a posthole describing the operative decisions regarding its exploration (F279-LZ).

Finally, Tarbat sheet also explain when operative and recording decisions had interpretative consequences. For instance, when the action of the diggers could affect the reliability of data: "*probably a posthole, oval shape exaggerated by overcut at northern end*" (F234-JGL). Similar situations could arise in the exploration of negative features when a layer was exclusively observed in section but not in plan: "*F24 was removed with 1103 and the underlying C1163 was only recognized in the first section, C1211 therefore only partially recorded in hachure plan*" (C1211-JB). A similar example was the initial definition of thin sand layer in plan but later confirmed in section: "*sectioned by fine trowelling which proved me deposit to be approximately 0.025m in depth, thus it was decided not to drawn or photograph*" (RJ-C1455). This example was very interesting due the significance of this layer for the interpretation of the feature (Example 5.10). Overall, these additional circumstantial and decision data were classified as instances of best practice because these testimonies provide a relevant informative background of observational and technical actions which frequently contributes to contextualize the credibility of data (Graphic 5.3).

5.5.2 Context descriptions

One of the most distinctive aspects of the feature system is the absence of cut sheets, thereby layers and deposits are the main entities being described in context sheets. This doesn't imply the absence of cuts information, but it was included as part of feature descriptions, and the consequences of this procedure shall be explained later. Likewise, some feature sheets included layer information, which was coded as such only if that information had not been previously recorded in a context sheets (Graphic 5.2). Generally, layer descriptions are recorded in context sheets, specifically in two sections. The first was a free text area called '*Description*' and the second a more standardized section labelled '*Components of visual estimates*' (Figure 5.1). Free-text descriptions usually provide a synthetic account of layers based in its commonest attributes:

> *C1360 was allocated to a very dark grey sand silt, frequents inclusions of large cobbles and charcoal flecks. Also, frequent inclusions of winkle in good to variable condition towards the end of the context. Homogenous with little colour variations. Fells very sticky and plasticity* [sic]. *Finds from this context include animal bone (10-50mm), slag and pottery. Large cobbles, up to 250mm seem to be concentrated towards the bottom of the deposit" (C1360-KJA/DW)*

In contrast, *Visual Estimates* supply a more analytic account of attributes (Figure 5.7). Some descriptive information is repeated in both areas, but *Visual Estimates* includes quantitative data about the distribution of components, whereas free-text description can be very useful to capture qualitative aspects that would be difficult to describe as visual estimates, for instance the concentration of cobbles and finds in the bottom of C1360 as in the example above.

Components from Visual Estimates

%	Distribution	Compaction	Structure	Purity	Colour	Size (mm)	Material	A/R/M	D/S/K
85		STICKY	CRUMBS	CLEAN	5YR 2.5/1	>5	SANDSILT		D
2	DISORDERED	SOFT	FLECKS+FRAGS	BURNT	BLACK	5-20	CHARCOAL		D
3	DISORDERED	FRIABLE	FRAGMENTS	VARIABLE	WHITE	5-15	SHELL		D
	DISORDERED	HARD	FRAGMENTS	VARIABLE	VARIABLE	20-50	SLAG		K
	DISORDERED	FRIABLE	FRAGMENTS	VARIABLE	VARIABLE	10-500	BONE		K
	DISORDERED						POTTERY		K

Figure 5.7. The description of a layer as visual estimates. (C1360-KJA/DW).

Likewise, Tarbat's sample was a rich source of multiple descriptive strategies. One of the most common was some sort of abstract account based in the idealization of attributes. "*C1490 allocated to slightly dark grey sand. Backfill of F180. Frequent charcoal fragments, flecks and fragments of animal bone. Rare inclusions of slag and rounded pebble*" (C1490-CKH). A less common descriptive pattern was a more 'exhaustive description' based in a more detailed account of attributes. For instance, the description of C1360 highlights the presence of cobbles towards the bottom of the context. These descriptions are particularly relevant when sch 'detailed observations' have interpretative or reflexive implications like in the following example: "*The* *upper deposit of context 1090 was more slag filled and the bottom more charcoal/ash concentration. This may have been the result of in situ burning prior to the removal of the post*" (F29-ST).

Tarbat sheets also provide clear testimonies of alternative processes of layer definition, for example when one context is defined by lumping various layers like in the example below. Generally, sheets do not include an explanation of the rationale of these decisions, but descriptive data provides crucial elements to understand this decision. For instance, C1412's description indicates that context definition is based in technical constrains to dissect each layer given their thickness. More important is noticing that such interpretative decision does not seem to compromise the credibility of data because this interpretation still captures the main attributes of layers and doesn't constrain the interpretation of formation processes.

<u>Example 5.1</u>

> *Upon excavation, C1412 proved to be between 15 and 60mm thick. Lying on a flat sandstone slab, which constituted the base of the hearth. This material had a laminated appearance comprising of alternating bands of charcoal, ashy silt and brown silt indicating the deposit was result of several burning episodes (C1412-ST)*

Altogether, these examples represent different standards to approach layer definition and description. Yet, such variability does not compromise the quality of records, instead they contribute to maintain it. Once more, although the architecture of the system imposes some rules, the system is flexible enough to capture different interpretative and descriptive strategies. Then these examples provide more evidence against the belief that standardized formats impose a strong homogenization of data and obscure interpretative decisions.

As explained before, layer descriptions require to indicate a type of layer. This information could be indicated within descriptions but it was also systematically recorded in the section '*identified as*' mentioning if a layer corresponded to a natural, dump, burial, makeup, surface layer and so forth (Figure 5.8). Then, when this task was accomplished, context sheets were classified as valid interpretations. On the contrary, unreliable sources normally include a description without stablishing the type of layer being described. But in some cases, unreliable documents only include a context number with no descriptive information at all (Graphic 5.3). In the process of analysis, it was noticed that some context sheets included a description but no visual estimates, whilst other sheets showed a misplacement of information. For example, it was recurrent that some sheets indicated ‘dump’ for the profile shape of a layer (Figure 5.8). These examples were counted as recording mistakes or minor failures without compromising the reliability of descriptions.

Examples of layer reflexivity are very similar to those observed in former projects. Persistent reflexivity was the most frequent form of layer reflexivity and usually describes cases of ambiguous boundaries but without compromising interpretative data. For instance, when describing the fill of a drain structure, a sheet mentions: “*Very difficult to determine the edges of this context. Forms the primary backfill of shallow drain F378*” (C1932-CKH). A notable characteristic of this project was the lack of sketches to represent ambiguous borders like in MRG95. Evidential testimonies usually described interpretive doubts affecting layer definition like in the following example that describes the lowest fill of a feature: “*Edges were difficult to define in places particularly on western edge, due similarities of material. Possible that cobbles in this context lined or were placed in the bottom for drainage. Cobbles were quite disorganized though*” (C1360-KJA/DW) (Graphic 5.4).

Finally, some context sheets included brief descriptions of the uncertainties and reconsiderations in the process of layer definition, therefore being very similar to the narrative of discovery in reflexive diaries. However, these accounts were normally recorded as introductory to descriptions: “*The edges of C2427 were initially impossible to define as they were covered by C2424. Once this layer was removed excavation of C2427 was possible*” (C2427-KM). However, the sample also included some examples of observational fluidity recorded with brief updated notes like in MRG95’s sheets. For instance, the initial account of a layer mentions uncertainty about its extension: “*C1535 also appears to continue across module D3 and under eastern section of INT24*” (C1535-DW). However, two years later when the exploration continued, uncertainty disappeared, and the description was updated: “*Context continued under baulk at a similar width for approx. 0.6m shallowing to a curved end*” (C1535-DW). Hence, these testimonies were coded as instances of observational reflexivity.

Y2.0	CONTEXT RECORD	Site: TR'04	Context: 2547

Intervention: 24 | Quadrant: D2

Easting: 881.69 | Northing: 991.85 | Height: 13.68 m AOD

Recovery level: A B C (D) E

Shape - In plan: SUBCIRCULAR | In profile: DUMPS.

Description: C2547 WAS ALLOCATED TO THREE SMALL DEPOSITS OF FIRED CLAY, LOCATED TO THE SOUTH AND SOUTHEAST OF HEARTH F148. MAXIMUM DIMENSIONS OF 20cm x 15cm, 10 x 4cm AND 10cm x 5cm, EACH COMPRISED A HARD DEPOSIT OF REDDISH BROWN FIRED CLAY, MOTTLED WHERE CRACKS HAVE BECOME FILLED WITH OVERLYING BROWN DEPOSITS. THE NORTHEASTERNMOST (AND LARGEST) DEPOSIT WAS SEEN TO ABUT THE MAKEUP STONES OF HEARTH F148, AND CONTAINED FRAGMENTS OF FUEL ASH SLAG AND A CRUCIBLE FRAGMENT.

Condition on recording: DAMP | M.C: REDDISH BROWN | Sieving regime (%): N/A

Components from Visual Estimates

%	Distribution	Compaction	Structure	Purity	Colour	Size	Material	A/R/M	D/S/K
97		HARD	FRAGS	FIRED	5YR4/4	VARIABLE	CLAY		K.
2	DISORDERED	HARD	FRAGS	BURNT	VARIABLE	<50mm	CRUCIBLE FRAGS		K
1	DISORDERED	FRIABLE	FRAGS	BURNT	VARIABLE	5-10	FUEL ASH SLAG		K
						2 - 6	Gravel		
						6 - 20	Gravel		
						20 - 60	Pebbles		
						60 - 200	Pebbles		
						200 - 600	Cobbles		
						> 600	Cobbles		

Samples: 100%. SAMPLED.

Stratigraphy:
Over: | Under: 1284?
Abuts: C1413. | Abutted by:
Same as: | Cut by:

Comments: FABRIC OF C2547 APPEARS THE SAME AS THE FIRED CLAY MAKEUP OF POSS FURNACE BASE F853 AND F479. ACTIVITY WOULD APPEAR TO BE CENTRED AROUND THE HEARTH F148

Belongs to feature: 148 | Identified as: FIRED CLAY DUMPS | Status: P / (S)

Date recorded: 20.07.04 | Recorded by: NJT.

FIELD ARCHAEOLOGY SPECIALISTS LTD — FAS

Figure 5.8. Context description for a dump of fired clay (C2547) related to a hearthe (F148). (C2547-NJT).

5.5.3 Finds and samples

Finds information was predominantly recorded in context sheets within layers descriptions, occasionally this information was included in feature sheets but always related to a specific context (Graphic 5.2). Additionally, visual estimates also indicate when finds had been kept (K) or discarded (Figure 5.7). One half of the context sheets didn't include finds information but unlike MOLA's sheets, Tarbat's format lacks a section to indicate finds absence and only three sheets explicitly indicate this with short notes: "*no finds were recovered*" (C2189-LOB). Then, the number of Tarbat sheets with no finds information was classified as unreliable testimonies (Graphic 5.3), although this mistakes are partially caused by the architecture of the format (Figure 5.1). No reflexive data related to finds collection was observed in Tarbat's records.

Sampling information was systematically recorded in context sheets with a short description of the type of sample, likewise the final column of visual components usually remarked which elements had been sampled (*S)* (Figure 5.8). In the opposite case, some sheets included short comment as "*none*" or "*none taken*", but the most frequently the space was cancelled, therefore context sheets provide a consistent record of sampling activity. Only the positive cases were counted (Graphic 5.2). Occasionally sampling decisions were explained indicating a clear purpose: "*an early medieval timber sent to Edinburgh for attempted dendrochronology*" (C2252-NJT). Alternatively, some sheets explained the reasons of collection: "*because of the large volume of iron slag present, a sample of half large sack was taken, mostly large pieces and the rest was discarded*" (C1306-IAJ).

On the contrary, when sheets only provided a brief description of the type of sample but no explanatory comments. For instance, a "*baulk sample, approximately 10L taken for flotation*" (C1348-RB) or "*1 bulk sample*" (C1412-ST), sampling decisions became evident after considering interpretative information: "*C1412 represents use of hearth F148*". Many of these cases included rich layers as charcoal/organic fills within fire features or shell assemblages, that had provided many clues about depositional processes and then deserved further investigation offsite. Likewise, another route for sampling action was when layers hadn't been properly understood onsite because evidence was not very telling. This included multiple examples of fills within some large ditches that in broad terms could be described as 'sterile' (Example 5.13).

There was only a small number of sheets in which sampling decisions were unclear due the lack of interpretative data and therefore were classified as unreliable records (Graphic 5.3).

In broad terms, records from Tarbat and MRG95 reflect similar principles for sampling choices. In general, fieldworkers decided to sample a layer that had demonstrated interpretative potential for having rich assemblages or when layers had been difficult to investigate onsite. The only difference between is that the sampling policy of MRG95 was more rigid because it demanded the systematic sampling of every charcoal/organic fill inside kilns. On the contrary, Tarbat's strategy was more selective as not every fill of hearth was sampled.

5.5.4 Other low-level claims

As in former case studies, low-level contrasts, expected absences and indicative absences were put together in a category due its low frequency. These claims were found in contexts and feature sheets alike (Graphic 5.2). As in former case studies, these claims describe various types of observations, however the most interesting examples contribute usually work as negative evidence: "*Due its layered nature and the lack of stones or other finds, it probably formed naturally*" (C1167-JB). Likewise, some of these testimonies positively contribute for the evaluation of a hypothesis as in the example below when a former interpretation is reformulated after considering the absence of expected elements: "*Burnt fill within F15 caused by the process of metalworking... It contains no slag and no bronze fragments so it may just be burnt rather than metalworking layer" (C1444-TJS).*

5.5.5 Stratigraphic information

Stratigraphic information was equally distributed in context and feature sheets being layer contacts the most frequent (Graphic 5.2). Specifically, context sheets focus in the description of layer contacts among deposits (*over*, *under*, *abuts*, *abutted*) and this format only includes one entry for cutting relations (*cut by*) (Figure 5.8). In contrast, feature sheets include entries for layer contacts (*overlies*, *abuts*, *abutted by*) and cutting relations (*cuts* and *cut by*) (Figure 5.11). Initially, this format seemed strange because physical relationships are among contexts, not features. Yet, the system made more sense after noticing that cuts are recorded within feature descriptions. Nevertheless, one also must be aware that interpreting the sequence stills requires further analysis of site records because Tarbat's formats don't include a section for recording stratigraphic relations, namely the Harris matrix. Only a very small portion of sheets included stratigraphic diagrams to represent the stratigraphic relations of specific layers defining a feature.

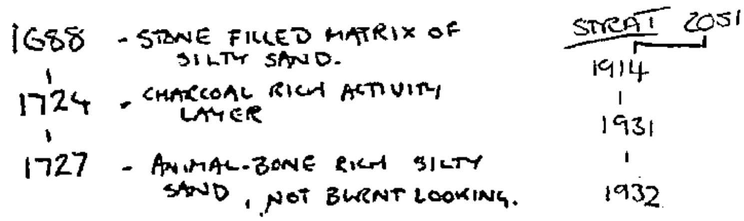
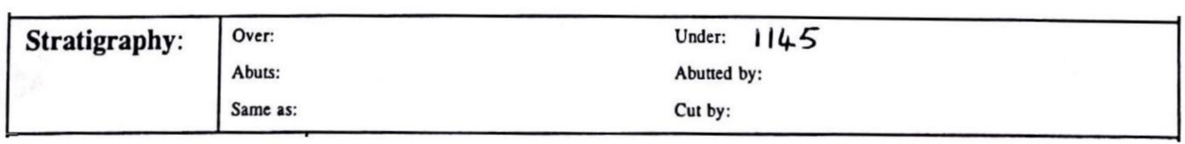


Figure 5.9. Stratigraphic diagrams in feature sheets. F299-JMM (left) and F378-CKH (right).

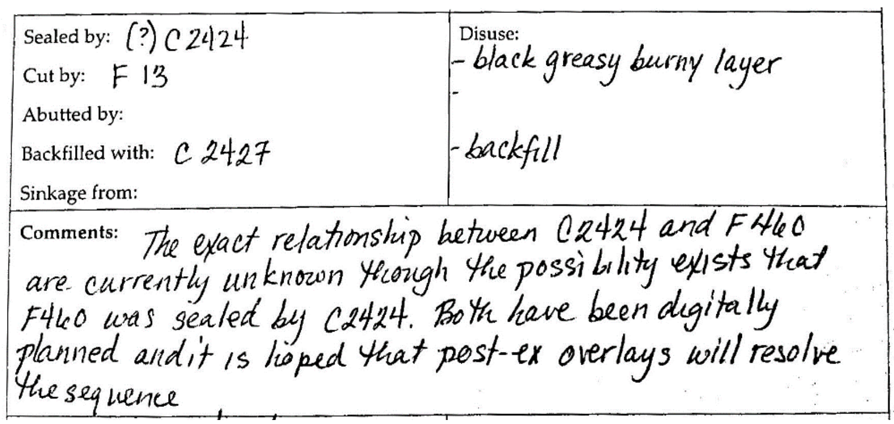
Almost one third of the context sheets recorded only one-layer contact (Figure 5.10). Initially, this looked like an incomplete record if one considers that stratigraphic units at least must have two stratigraphic relations with an earlier (usually underlying) and a later unit (usually overlying). However, since Tarbat's sheets do not represent stratigraphic relations, then such documents are not necessarily incomplete. Likewise, some sheets recorded three or more layers contacts, however it's unclear whether all of them are relevant for interpreting the sequence. Overall, Tarbat's system show less emphasis in stratigraphy, which is reasonable if one considers that the sequence is relatively simpler than in London. Around 10% of the context sheets lacked LC information and these were classified as unreliable sources (Graphic 5.3).



Stratigraphy:	Over:	Under: 1145
	Abuts:	Abutted by:
	Same as:	Cut by:

Figure 5.10. The stratigraphic section of a context sheet (C1146-TJS).

Uncertainty and fluidity in layer contacts are the commonest forms of stratigraphic reflexivity at Tarbat. Once more, many uncertainty cases are symbolically recorded by adding a question mark next to a specific problematic contact (Figure 5.11). Occasionally, some sheets included a textual note describing the symbolic annotation: "*The exact relationship between C2424 and F46 are currently unknown thought the possibilities exist that F460 was sealed by C424*" (KM-F460). In theory, all these doubts should be relevant when interpreting the sequence and hence they were classified as instances of evidential reflexivity (Graphic 5.4). However, in most cases symbolic and textual information was insufficient to understand the cause and implications of doubt, therefore being classified as problematic evidential testimonies. There are only a few examples that briefly described the implications of ambiguity in layer contacts for stratigraphic relations: "*It is believed that F386 was cut at the same level as F383. There is however no clear stratigraphic relationship between these features*" (LG-386). However, since no sequence claims are included in the first place, it is probably unreasonable to expect anything else than a brief description of doubts. Fluidity was coded from corrections made to layer contact and where classified as instances of observational reflexivity but anything else couldn't be noticed.



Sealed by: (?) C 2424 Cut by: F 13 Abutted by: Backfilled with: C 2427 Sinkage from:	Disuse: - black greasy burny layer - - backfill

Comments: The exact relationship between C2424 and F460 are currently unknown though the possibility exists that F460 was sealed by C2424. Both have been digitally planned and it is hoped that post-ex overlays will resolve the sequence

Figure 5.11. A fragment of F460 showing stratigraphic information. (F460-KM).

5.5.6 Correlations

The frequency of correlation claims is low and this information was recorded in the stratigraphic sections of context and feature sheets (Graphic 5.2). In general, these claims were recorded using the formula 'same as' which is equivalent to the symbolic code (x=y). The main distinction in the feature system is that context sheets include correlations among layers: "*same as: 1094/1093*" (C1143-TJS), whereas feature sheets correlate cuts: "*same as: F9*0" (F52-JB). Tarbat's system also uses the notion of correlation in two senses, either to represent discontinuous layers or features once being connected, but also to link two contexts/features that belong to the same entity but had been dug in two interventions. In the second case, sheets normally included explicative notes or sketches to clarify this meaning. For instance, sheet for F23 mentions "*same as F99*" (F23-RTJ/ST) but a note and a sketch on the backside explain: "*F23 was only partially excavated in the TR96 season. In the edge of F2 the remainder of this gully could be identified, cut through F90 (The continuation of F52. Another gully to the south of*

F2)" (F23-RTJ/ST) (Figure 5.12). However, the same note and drawing explain the correlation between F52 and F90 that represents two stratigraphic units separated by F2. In general, correlation claims were classified as valid testimonies because records usually provide enough information to understand the recording process. There is a small number of sheets that couldn't be evaluated because only one of the related sheets was collected (Graphic 5.3).

A few cases of uncertainty situations were recorded adding a question mark next to correlation claims. These instances where always associated to the stratigraphic meaning of the concept and were classified as instances of relevant reflexivity, but unlike MOLA's sheets, Tarbat records never included explicative comments hence being classified as instances of problematic reflexive testimonies (RP) (Graphic 5.4).

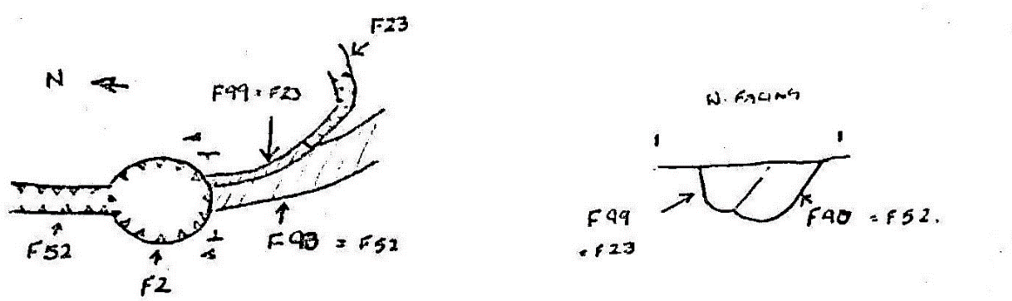


Figure 5.12. Sketch illustrating two types of correlations. One for F23=F99 excavated in two interventions and another for F52=F90 separated by F2. (F23-RTJ/ST).

5.5.7 Assemblages

Tarbat's sample described two types of groups of finds. The first included clusters defining stratigraphic units like in the following example: "*dump of shells identified after removal of C1427, made up of winkles and lumpets. Thickness of layer less than 2cm*" (C1428-LG). Whilst, the second type included groups of finds defining stratigraphic units in which the spatial distribution of finds was crucial evidence and records include this information, therefore being classified as assemblages. TRS2 included a few examples of burial assemblages (Graphic 5.2) but one of the most interesting was the description of an animal skeleton related to F304.

> *"C1734 was allocated to a partially articulated, mature female cow skeleton. During definition, it became apparent that the skeleton had been crudely butchered (dismembered) before deposition. All four limbs, skull, ribcage and vertebrae had been separated…Obvious butchery marks on pelvis and*

therefore sacrum, and rib facets butchered and still articulated with vertebrae. Maturation of bones (fused epiphyses and radius) would suggest old dairy cow therefore not suitable for meat hence full burial? Absence of tail is a mystery" (C1734-CAS)

As one can observe in the fragment above, the interpretation of the depositional process of the skeleton was derived from the position and the attributes of bones as both provide crucial evidence to interpret its burial process. Due the short number of assemblage information, no validity analysis was made, although the former example represents an instance of reliable testimony. For the same reason, no instances of assemblage reflexivity were identified.

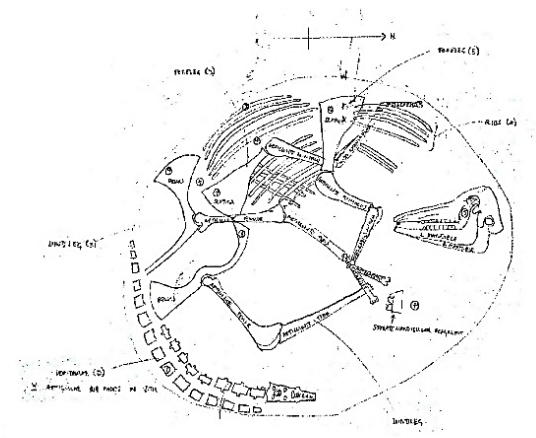


Figure 5.13. Sketch of a cow burial showing the spatial arrangement of bones (C1734-CAS).

5.5.8 Structural elements

Almost a third of context sheets describe structural elements such as floors, linings, packing stones or makeup fills whereas two thirds of the feature sheets describe structural features structures such as hearths, flues, wells or postholes (Graphic 5.2). As in former projects, records describe two types of epistemic situations. One when the structural character of a remain is evident and then context and feature sheets only have to describe the things being observed. For example, the fragment below describes a stone frame from hearth.

Example 5.2

C1413 comprises a four angular sandstone blocks (length between 200x600mm, and width 150mm, depth 100mm. Set in square arrangement ending in a single flat sheet of sandstone […] The upper surface of the stone appears smooth and worn. The hearth base is fragmented and burn" (C1413-ST)

Subsequently, feature sheets provide a feature description that summarizes information from structural layers (Example 5.2) and non-structural layers (Example 5.1). In consequence, the feature description repeats some information already contained in context sheets like the fragment below that resumes the description of a stone hearth (Figure 5.15). These examples are identical to the description of pottery kilns at MRG95 and almost every structural component at B49.

Example 5.3

"F148 appeared to consist of an arrangement of angular sandstone blocks set on edge creating three sides of a square approximately 0.8m across. Within this formation –C1412-a brightly heavily mottled deposit of reddish yellow silt clay ash sat on a fragmented sandstone slab. This material was excavated and appeared to be the remains of primary burning within the hearth of the structure …" (F148-ST/NJT)

On the contrary, the second type of interpretative situation covers cases when the structural character of a remain is not obvious because evidence is fragmentary or marginal. Unlike MRG95 that had more frequent examples of fragmentary evidence, Tarbat had more recurrent examples of marginal evidence like the following testimony.

Example 5.4

Backfill of F460, a posthole. C2427 was a dark grey sandy silt with inclusions of charcoal and gravel as well as lenses of light brown sand and bits of yellow clay. A central block of slightly darker brown sandy silt without gravel was evidence for a post pipe that could be seen in section. (C2427-KM).

Tarbat sample provided numerous examples that illustrate the use of different types of marginal evidence to interpret the structural character of remains, including the physical contacts among layers as in the following example: "*C2242 appeared to form a wedge of gravel underneath the southern edge of sandstone slabs C2243, where the ground levels dips south, so possible used to stabilize a level hardstanding F434*" (C242-CAS). This sample confirmed the low proportion of unreliable structural information, only a few sheets failed to explain the interpretive process of non-evident structures (Graphic 5.3). These cases include a few sheets describing some surfaces related to the main road F469, in which the connection between descriptive and interpretative data is not very clear: "*reddish brown clay silt burrowed with darker reddish patches (very decayed sandstone) virtually stains*" (C3206-DMK). However, these unreliable interpretations of structural layers didn't seem to affect the credibility of the structural interpretation of the feature because related context sheets provided supporting data. Specifically, some sheets explained the how the physical relations among a group of extensive layers suggested the presence of a road (Figure 5.14). Nevertheless, there could be cases in which the unreliability of a layer interpretation could affect the credibility of a feature claim.

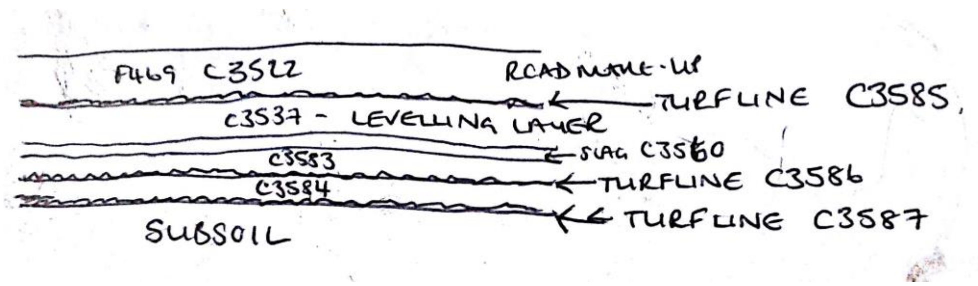


Figure 5.14. Sketch of a section for road F469. (C3522-LAC)

Tarbat records also include numerous examples of structural reflexivity. Some of these testimonies portray the interpretative doubts and reconsideration in the process of discovery of structural layers and features. Generally, this information is included in the introductory lines of descriptions as in the following example: "*F353 was identified in 1999 as a possible posthole, but further definition in 2004 revealed a sub circular feature. Possibly a furnace base comprising a ring of fire clay (2540) containing charcoal rich burning deposit (C1815)*" (F353-NJT). One of the most interesting examples of observational reflexivity was a testimony of the discovery of a posthole (F460), this example clearly illustrated the irrelevance of the irrelevance of this information to stablish the credibility of interpretative data (below), contrary to explanatory account that describes the interpretive process of marginal evidence, therefore being more relevant for stablishing the credibility of data (Example 5.4).

Example 5.5

"Backfilled posthole which could be seen clearly in the section of the wall of F15which cuts through it. The top edges were much harder to define as they were partially covered by C2424. Once C2424 was removed a second layer (C2428) complicated definition. C2428 was cleaned and sectioned at which point it was determined that it was out of sequence with F40 and should be left. At this point the edges of F460 became clear and I was able to continue excavating. F460 proved to be a backfilled posthole with clear evidence of postpipe" (KM-F460).

Tarbat sheets also include some very few examples of updates and corrections in feature descriptions when the excavation of a feature had lasted for several seasons, thereby adopting a similar strategy to MRG95. However, any of these cases implied an interpretative reconsideration. For instance, the fragment below is an update to the initial description of F148 annotated in the backside of the sheet: "*Further definition in 2004 revealed that a horizontal red sandstone slag (2541) was part of F148 and appears to form an original lining of the base of the hearth*" (F148-ST/NJT). These types of examples were considered instances of observational reflexivity, yet they didn't seem irrelevant like chronicles of discovery because updated information contributes to the accuracy of descriptions.

Additionally, the sample includes a few cases of persistent reflexivity that usually describe situations when the structural character of something has been defined, but fragmentariness produces some doubts regarding some attributes, most often the extension of structures as in the following example of a stoned-lined hearth: "*F495 was severely truncated at its eastern side by F13 and its construction at this point is unclear*" (F495-NJT). Cases of evidential reflexivity are the most frequent for structural information (Graphic 5.4). Some of these testimonies describe situations when the structural character of remains was unclear and then sheets include shorts discussions as in the example below (Figure 5.12).

<u>Example 5.6</u>

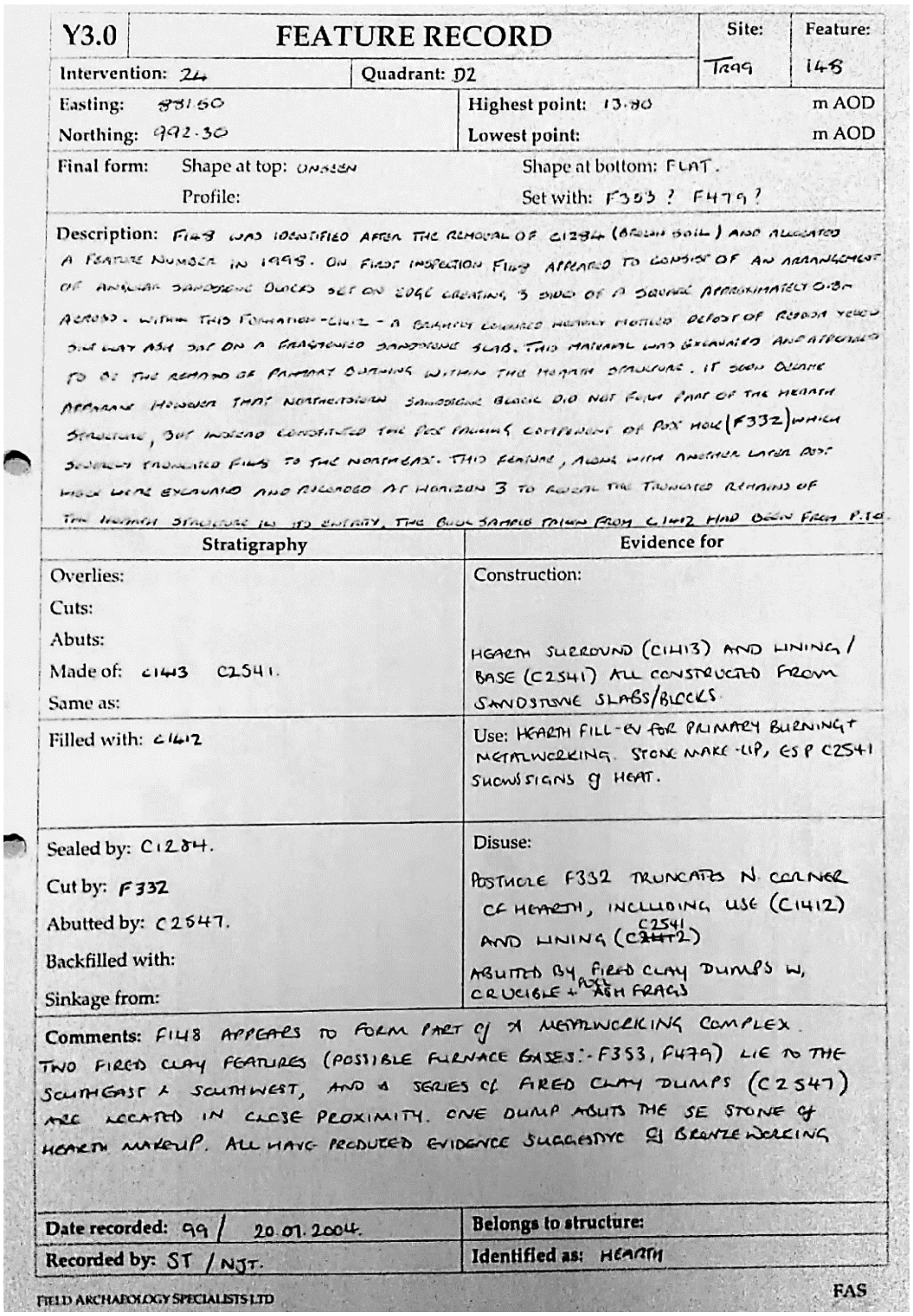
> *"F52 is of uncertain use. No evidence of postholes or stakeholes were found; as the feature could have hold a fence. It could also have been a drainage channel, there was however no evidence of naturally built up silting in the base of the cut" (F52-JB)*

Another type of reflexive situation was when the structural character of remains was clear, but the type of structure was ambiguous. Occasionally, some of these testimonies could even consider spatial signs when discussing puzzling evidence and interpretative possibilities:

<u>Example 5.7</u>

> *"Shape and arrangement of the stones implies use as the floor of a building or a pathway where alignments respects road F18. More likely to represent a pathway associated with the road, since it slopes down from N-S, like the road, rather than the floor of a building which would probably have been levelled" (F16-CKH/ST).*

Nevertheless, there is an important number of problematic evidential testimonies (Graphic 5.4). As in previous cases, most of these testimonies indicate two or more interpretative possibilities but without developing an argumentative discussion as in the following example: "*Alignment and presence of a large cobbles in the bottom of C1360 suggest field drainage. Other possibility could be cut for another boundary wall foundation*" (F134-KJA/DW).



Y3.0 — FEATURE RECORD | Site: TR99 | Feature: 148

Intervention: 24 | Quadrant: D2

Easting: 88150 | Highest point: 13.80 m AOD

Northing: 992.30 | Lowest point: m AOD

Final form: Shape at top: UNSEEN | Shape at bottom: FLAT.

Profile: | Set with: F353 ? F479 ?

Description: F148 WAS IDENTIFIED AFTER THE REMOVAL OF C1284 (BROWN SOIL) AND ALLOCATED A FEATURE NUMBER IN 1998. ON FIRST INSPECTION F148 APPEARED TO CONSIST OF AN ARRANGEMENT OF ANGULAR SANDSTONE BLOCKS SET ON EDGE CREATING 3 SIDES OF A SQUARE APPROXIMATELY 0.8m ACROSS. WITHIN THIS FORMATION - C1412 - A BRIGHTLY COLOURED HEAVILY MOTTLED DEPOSIT OF REDDISH YELLOW [illegible] ASH SAT ON A FRAGMENTED SANDSTONE SLAB. THIS MATERIAL WAS EXCAVATED AND APPEARED TO BE THE REMAINS OF PRIMARY BURNING WITHIN THE HEARTH STRUCTURE. IT SOON BECAME APPARENT HOWEVER THAT NORTHERNMOST SANDSTONE BLOCK DID NOT FORM PART OF THE HEARTH STRUCTURE, BUT INSTEAD CONSTITUTED THE POST PACKING COMPONENT OF POST HOLE (F332) WHICH SEVERELY TRUNCATED F148 TO THE NORTHEAST. THIS FEATURE, ALONG WITH ANOTHER LATER POST HOLE WERE EXCAVATED AND RECORDED AT HORIZON 3 TO REVEAL THE TRUNCATED REMAINS OF THE HEARTH STRUCTURE IN ITS ENTIRETY. THE BULK SAMPLE TAKEN FROM C1412 HAD BEEN FROM P.T.O.

Stratigraphy	Evidence for
Overlies: Cuts: Abuts: Made of: C1413 C2541. Same as:	Construction: HEARTH SURROUND (C1413) AND LINING / BASE (C2541) ALL CONSTRUCTED FROM SANDSTONE SLABS/BLOCKS.
Filled with: C1412	Use: HEARTH FILL - EV FOR PRIMARY BURNING + METALWORKING. STONE MAKE-UP, ESP C2541 SHOWS SIGNS OF HEAT.
Sealed by: C1284. Cut by: F332 Abutted by: C2547. Backfilled with: Sinkage from:	Disuse: POSTHOLE F332 TRUNCATES N. CORNER OF HEARTH, INCLUDING USE (C1412) AND LINING (~~C2412~~ C2541) ABUTTED BY FIRED CLAY DUMPS W, CRUCIBLE + ASH FRAGS

Comments: F148 APPEARS TO FORM PART OF A METALWORKING COMPLEX. TWO FIRED CLAY FEATURES (POSSIBLE FURNACE BASES? - F353, F479) LIE TO THE SOUTHEAST & SOUTHWEST, AND A SERIES OF FIRED CLAY DUMPS (C2547) ARE LOCATED IN CLOSE PROXIMITY. ONE DUMP ABUTS THE SE STONE OF HEARTH MAKEUP. ALL HAVE PRODUCED EVIDENCE SUGGESTIVE OF BRONZE WORKING

Date recorded: 99 / 20.01.2004. | Belongs to structure:

Recorded by: ST / NJT. | Identified as: HEARTH

FIELD ARCHAEOLOGY SPECIALISTS LTD — FAS

Figure 5.15. Feature sheet of F148. The feature was initially described by ST in 1999, but the description continued and concluded in 2004 by NJT.

5.5.9 Features

Feature claims were the most frequent type of information, but this shouldn't be a surprise considering the characteristics of the sample (Graphic 5.2). Moreover, feature information appears in context and feature sheets, although the type of information is different. Context sheets only include a reference to a feature number in the section '*belongs to feature*' (Figure 5.8), whereas feature sheets include a feature description that summarize the information from layer descriptions and also include cuts information in the case of negative features.

<u>Example 5.8</u>

> *"F378 is a covered drain running across mod A5 and up to LOE. Only the area of collapsed lid was excavated allowing the view of accumulation within. F378 is constructed of large sandstone slabs lying over shallow (0.10m gully). Presumably used to aid drainage over this area. Approx. 3m in length up to the section. F378 appears to curve southwards before edge of excavation…" (F378-CKH).*

However, this standard was not always achieved, and many descriptions of negative features were reduced to cut descriptions: "*F68 was first defined after the removal of C1040/C1025. This feature is step sided hole approximately 19cm in diameter and 26cm deep. The edges of this feature were clearly defined at its base but less clear at the surface*" (F68-EJH). Occasionally, positive features had descriptive problems too because testimonies provided a physical description of remains but without giving an account of the group of layers, like in the following example: "*First identified after the removal of C1284 as a stone-built wall in quad D. The feature ran from the east facing section of MODD3 and curved in a SW direction through D2 and D1 towards the 'millpond' area*" (F149-CAS). Around one half of the feature descriptions showed these kinds of problems and were classified as recording mistakes (Graphic 5.3). This made clear that the real difference between single context recording and the feature system was not interpretative. More exactly, the difference lies in the type of descriptions provided by each method because the purpose of a feature sheet is to provide synthetic account of groups. Yet, when this goal is not achieved, then there is no real difference with single context recording that works linking different pieces of data distributed across multiple context sheets.

A fundamental requirement of feature descriptions is stablishing its type. For instance, a hearth, posthole, shell-midden, rubbish-pit and so forth. In broad terms, features types can be classified in two classes, namely structural and non-structural. This distinction was positively executed in 70% of the sample, whereas in 17% of the cases, sheets reported uncertainty. Altogether, these cases were considered valid feature information (Graphic 5.3). In general, the type of feature was indicated in descriptions but this was confirmed in a small field in the lower part of feature sheets, labelled '*identified as*' (Figure 5.15). In contrast, negative features were commonly described as pits or ditches. As mentioned before, these categories are problematic because a pit can be structural if used for storage or non-structural if used for dumping. Hence, being classified as instances of unreliable feature information. Likewise, testimonies of negative features that didn't include masonry evidence usually failed to consider cut's evidence to interpret its type. For instance, the following description probably corresponds to some form of non-structural feature due its irregular cut shape, but this is merely described as a shallow-scoop which is unclear to indicate its type: "*Upon excavation F278 was resolved as a shallow scoop well defined against surrounding clay silt deposit with a slightly irregular base and sloping edges (0.15m deep sides and sloping at 20°*" (F278-ST). This was an important difference with MRG95 in which diggers had more awareness of the relevance of cut's evidence.

Additionally, some sheets showed another form of recording problem. Specifically, there are some features that only group one context sheet, and this was conceptually inconsistent because a feature should group two or more contexts. Many of those anomalous groups are consequence of Tarbat's recording strategy because context sheets only describe layers whereas feature sheets describe cuts and features altogether. Hence, some of these groups were legitimate. Nevertheless, there were some cases in which one layer was assigned a context and a feature number. These included various structural elements as stone surfaces and platforms (F16/C1118, F433/C2202, F474/C2496, F522/C3071, F525/C3136), walls (F394/C2024, F432/C2348) and one big dump (F109/C1306). However, as one compares the descriptive and interpretive content in context and feature sheets, this was practically the same. For instance, the layer description for C1118 already explains its structural character:

> *"A concentration of flat sandstones which in plan appear not as random scatter. Many are shattered, the majority being thin and squared. The alignment of the concentration on the western side appear to respect the line of the road F18.Although the thickness of the stones is variable, about 60% appear to be thin (approx. 10-40mm), which implies use as flagging" (C1118-CKH)*

Similarly, the feature description (F16) practically reiterates the structural interpretation of C1118 as ‘crazy paving’. However, even if the feature sheet includes spatial data, this certainly doesn’t justify the enlargement of the archive or treating C1118 as a group.

> *“Identified during excavation of 1099 as a concentration of flat thin sandstones mostly rectangular in shape although shattered into several pieces. Towards the south there is an area of fragments, which are set almost in the style of ‘crazy paving’. The alignment of the western edge runs N/S parallel to the edge of F18 (road) and the eastern and southern side continue outside the edge of the intervention (F16-CKH)*

Initially, it was considered that such inconsistencies responded to a recording strategy that consisted in describing every structural remains as a feature, independently whether these represented a group of layers. To test this idea a small sample of floating contexts (FC) was examined. However, these documents provided numerous examples of ungrouped structural layers and not defined as features. For instance: “*a spread if mixed pebbles approx. 1.25 m. long per 0.5m wide, tight against the south wall of structure 9 and the eastern end of F509 forming a possible gravel surface*” (FC-2243-RTJ). A similar case was the description of a structural surface very similar to F16/C1118: “*A concentration of large flat angular stones… The stones appear set together in plan, forming a possible paved area*” (FC1227-CKH). The last example is important for two reasons. First, it confirms an inconsistency in the recording strategy and secondly, because F16/C1118 and FC1227 were recorded by the same digger (CKH), then showing that participants could be inconsistent when defining features.

In sum, 12% of the feature descriptions are unreliable because they fail to define its type whereas 8% represent invalid groups formed by one deposit. Altogether these were defined as invalid feature claims which cover around 20% of the feature sheets. This number is not consistent with the proportion of reliable feature descriptions (87%), but this is consequence of the quantitative procedure because inconsistent groups were counted twice if these has been successful to define its type (Graphic 5.3). Finally, if one reconsiders these problems alongside recording mistakes in feature description, these indicate an important conceptual ambiguity in the status of features which sometimes are described as cut, sometimes as groups and sometimes as layers.

Feature reflexivity was the most frequent form of reflexive claim (Graphic 5.4). Some feature descriptions include testimonies of uncertainty and fluidity related to the process of clearance then being classified as instances of observational reflexivity as in the example below.

> *F172 was first identified as a spread of rubble (C1465) underlying C1444, with an ash black deposit overlying its northern limits. On removal of C1405, it became evident that this rubble constituted the final backfill of a pit feature. Further excavation revealed a stepped sided sublinear gully running NE/SW with rounded norther and southern ends" (F172-*CKH).

Examples of persistent reflexivity describe cases of ambiguity when defining the extent of a feature but without compromising its interpretation. For example, the fragment below describes the difficulties when defining the lower end of the cut from F378: "*very difficult to see the base of the cut due to invisibility to see edge*" (CKH-F378), but the available evidence provided a lot of support for the definition of a covered drain (Example 5.8). Finally, evidential testimonies describe three cases of interpretive doubts when defining features. First, there are cases when the type of feature was unclear, however these cases were very common for structural features (Example 5.6 and Example 5.7). Secondly, uncertainty could arise when the number of defined features was unclear, this was a recurrent problem derived from ambiguous layer contacts among cuts like in the example below (Figure 5.16):

<u>Example 5.9</u>

> *"F60 to the north of F24 could possibly have been the end of F52. It was so truncated by F24 and F37 that it was not possible to tell if it was related…F60 could either be the end of F52 or related to the shallow scoop in F52 (which could have been a feature in its own right. Although the fill appeared identical to that in F52..." (F52-JB/ST)*

Finally, feature reflexivity could emerge due grouping ambiguity, especially when it was unclear whether a layer belonged to a group. Most often, this ambiguity emerged when investigating negative features, especially when it was uncertain if a fill was contained within a cut: "*possible fill of F180 but dipping steeply down onto the loose fill of F70*" (C2738-RTJ). Only one example for positive features was found in the description a shell-midden (F156) defined by three superposed shell deposits (C1348, C1427 and C1428) and interrupted by two layers of silty sand (C1454 and C1455). Then, when a digger discovered an additional layer of shell below (C1467), he questioned whether this "c*ould be a continuation of shell–midden F156*" (C1455-RBM).

Example 5.10

> *"There is soil accumulation between bottom shell layer C1428 of F156 midden and new shell layer beneath this context which may imply a disuse, change of activity witnessed by C1456 and subsequent return to shell dumping area. But the accumulation as seen in E baulk of INT24 beneath F156 and as noticed during excavation is not great, certainly not more than 20-30mm so any disuse period is not for long. Leading me to believe that C1455 and C1456* [sic] *are related to the shell midden activity and not something entirely separate in between" (C1455-RBM).*

This example was very important to demonstrate the importance of including an argumentative discussion of evidence. This contrasts with the most recurrent evidential testimonies in context sheets that only include a brief description of an interpretive doubt: "*possibly a backfill of F527*" (NJT/C3154) or a question mark next to the associated feature number "*F471?*" (C2490-DW). Therefore, classified as instances of problematic reflexive testimonies. Less frequent are examples of interpretive reconsiderations, but testimonies always fail to discuss the evidence that motivated of such changes: "*originally thought to be the contexts in the pit F2, this feature number was allocated retrospectively when it became clear that the lower fills (now the context of F2) were of a different nature*" (F43-AR). In sum, one of the main problems with Tarbat records was the high amount of inadequate evidential testimonies for feature aspects (Graphic 5.4).

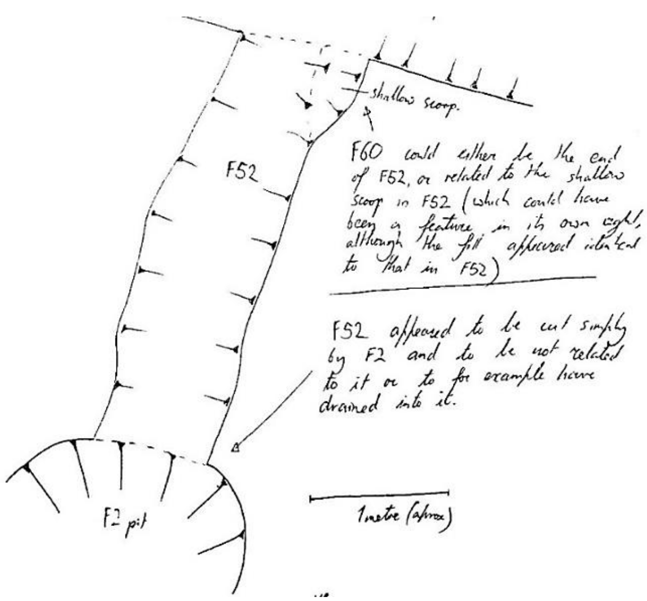


Figure 5.16. The sketch illustrates and discusses the unclear boundaries between F52 and F60 (F52-JB/ST).

In sum, the analysis of feature information has provided many examples of recording problems due the lack of clear standards for feature definition and recording. Similarly, reflexive information provided many elements to reconsider the recording strategy. For instance, it's very likely that the invested efforts in narratives of discovery should be better invested explaining and discussing situations of evidential reflexivity. Finally, another important aspect is noticing the similar diversity of interpretive and reflexive challenges at Tarbat and MRG95 due the higher diversity of feature-types which clearly contrasts with B49 where the type of remains was highly homogenous.

5.5.10 Formation processes

Another distinctive aspect of the feature system is its approach for recording formation processes. In the first place, context sheets must include an standardized box in the right bottom corner labelled '*Status*' where diggers should select one option, *P* for primary layers and *S* for secondary deposits (Figure 5.8). However, some sheets also included an explanatory account in the section for comments. Additionally, feature sheets include another section to record the depositional history of features whereby events will be interpreted as *Evidence for construction*, *use*, and *disuse* (Figure 5.15). Thereby, this is another area where there is a degree of repetition in the context of both types of sheets and this explains the close values in the frequency of formation process claims in context and feature sheets (Graphic 5.2). For example, the fragment below describes the depositional history of F148 that has been derived from the investigation of layers C1412 (Example 5.1) and C1413 (Example 5.2) among others (see also Figure 5.17).

> *"Construction: Heart surround (C1413) and lining base (C2541) all constructed from sandstone slab blocks. Use: Hearth fill event for primary burning metalworking. Stone makeup C2541 showing signs of heat. Disuse: Posthole truncates N corner of hearth" (F148-ST/NJT).*

Likewise, cuts can be interpreted as constructive evidence of negative features as in the example below, and truncations can be interpreted as evidence of post-depositional processes as in the example above. Then, the interpretation of cuts is an important difference in the depositional data of context and feature sheets in Tarbat's system.

> *"Construction: F378 appears to be a shallow gully cut to drain water or other material. Covered with large sandstone slabs (C2051). Appears that sides are not lined but stone blocks abut the cut doing similar function. Use:*

> *Drain/Covered culvert. Disuse: F378 appears to have two phases of silting/backfilling (C1931 and C1932). The lid C1914 had partially collapsed" (F378-CKH)*

Once more, the formation processes of masonry elements are recorded with great precision and context and feature sheets generally explain very well the evidence for construction, modification and destruction events. In contrast, the depositional data of negative features and sedimentary layers are problematic very frequently. For example, there is an important number of context sheets claiming the secondary character of many sedimentary by ticking the predetermined option (S). Besides, these sheets commonly describe such layers as backfills, but documents frequently provide no further explanation. Even if those claims are correct, the problem is not a lack of interpretation but a lack of support (Figure 6.2). Evidently, these represent cases of unreliable data which also affected the credibility of disuse data in feature sheets. Additionally, it was observed that records of negative features often fail to include constructive information because cuts are not considered as constructive evidence (Figure 5.18).

Y3.0	FEATURE RECORD	Site: TR99	Feature: 345
Intervention: 24	Quadrant: B0		

Easting: 371.30	Highest point: ~~1003.25~~ m AOD
Northing: 1003.25	Lowest point: m AOD

Final form: Shape at top: SUB-RECTANGULAR — Shape at bottom: IRREGULAR
Profile: U-SHAPED — Set with:

Description: F345 WAS IDENTIFIED AT HORIZON 3B AS A FAIRLY WELL-DEFINED ELONGATED FEATURE IN THE S E CORNER OF MODULE B0. THE BACKFILL OF THE FEATURE CONSISTED OF ~~[illegible]~~ 3 CONTEXTS OF DARK BROWN SILTY SAND (C1827, C1835, C1836) THE FEATURE MEASURES C. 250 CM IN LENGTH, C. 50 CM IN WIDTH AND C. 40 CM IN DEPTH. THE EDGES OF THE FEATURE WERE FAIRLY WELL-DEFINED AND EASY TO TRACE EXCEPT TO THE N-W, WHERE IT MAY ~~[illegible]~~ TRUNCATE ~~[illegible]~~ A LATER FEATURE.
F345 UPON EXCAVATION HAD A DISTINCTIVE U-SHAPED PROFILE WITH SIDES SLOPING ALONG ITS LENGTH AT BETWEEN 70 AND 85° BOTTOMING OUT ONTO A SLIGHTLY CONCAVE BASE. AT THE TERMINAL ENDS OF THIS ELONGATED PIT THE SLOPE WAS MUCH MORE GRADUAL (35-60°), DISTURBED BY OUTCROPPING FRAGMENTS OF STONE AND BONE FROM EARLIER DEPOSITS. F345 WAS EXCAVATED IN QUADRANT TO PRODUCE BOTH A LONGITUDINAL AND LATERAL SECTION. ALTHOUGH SOIL DIFFERENCES

Stratigraphy	Evidence for
Overlies: Cuts: F366 F561 BOTTOM 1801-1865 Abuts: Made of: Same as:	Construction:
Filled with:	Use:
Sealed by: C1621? Cut by: Abutted by: Backfilled with: C1827, C1835, C1836 Sinkage from:	Disuse: BACKFILLED WITH 3 DISTINCTIVE DEPOSITS CONTAINING A MIXTURE OF SLAG BONE AND SHELL IN VARIOUS QUANTITIES.

Comments: THE SHAPE AND PROFILE OF F345 ARE UNUSUAL AND DISTINCTIVE. NO EVIDENCE FOR FUNCTION, HOWEVER, COULD BE IDENTIFIED IN SECTION, PLAN OR BACKFILL. NO EVIDENCE FOR ANY STRUCTURAL COMPONENTS WERE LIKEWISE FOUND.

Date recorded: 25.08.99	Belongs to structure: —
Recorded by: LG	Identified as: PIT.

FIELD ARCHAEOLOGY SPECIALISTS LTD — FAS

Figure 5.18. Feature sheets of a pit with an unreliable information of disuse events and failing to interpret constructive and use events (F345-LG).

On the contrary, reliable depositional in context sheets always included a small explanatory account like in the following example that interprets a dark layer with larger quantities of slag and stone as a dumping event considering that: "*Some of the stones were stained by the matrix but they don't appear themselves as having been burn*" (C1006-JB). Likewise, another fill related to this feature was interpreted as a primary deposit considering the presence of slag and bronze fragments, in addition to: "*There was also a small amount of burned bone and flecks of charcoal*" (C1008-JB). Hence, the depositional history of the feature summarizes the evidence from primary fills, but the credibility of disuse event C1006 can be checked in its context sheet.

> *"Use: Probably used as a metal working smelting pit. Both fills contained concentrations of slag and waste bronze fragments. Traces of burning on the west edge of F4 provide evidence for burning in situ. The result of the first firing was C1008. C1008 was then cut through for another firing C1007. Disuse: Backfilled with C1006 which contained mainly stones (including two worked fragments if worked millstone and a large block), and large pieces of slag" (F4-JB)*

Still, there is an important problem with F4's testimony as it fails to consider cut's evidence for constructive aspects even-though the document notices burning traces in the cut. Initially, it was unclear whether this absence represented a real interpretative problem because in most cases, feature sheets included a cut description (Figure 5.18). Nevertheless, some sheets provided strong clues to confirm the unawareness of cuts as constructive evidence when diggers included a note that reads: "*No evidence for*" constructive processes (F52-JB/ST). However, some of these testimonies were corrected afterwards by someone else (Figure 5.19). Finally, there are some feature sheets that didn't include any information in 'Evidence for'. Many of these sheets are linked with contexts sheets that didn't include depositional data.

Stratigraphy:	Evidence for:
Overlies: Cuts: F60, F106, Abuts: Made of: Same as: F90	Construction: No evidence F52 CONTINUES TO THE SOUTH OF F2 AS F90 CURVING EASTWARDS AS IT RUNS SOUTH. IT CUTS F60 TO THE EAST AND [illegible] FEATURE F106. F90 CLEARLY CUTS ROAD F13 AND IT MUST BE ASSUMED THAT F52 IS A POST ROAD FEATURE ASWELL.
Filled with:	Use: No stake or post holes were

Figure 5.19. Section for constructive evidence of a feature sheet. (F52-JB/ST).

In general, four problems which affect the reliability of depositional data were identified. 1) Cases of unjustified depositional data in context sheets, 2) which affected the credibility of depositional data in feature sheets 3) limited interpretation of depositional data in feature sheets for failing to interpret cuts' evidence and 4) a lack of depositional data in features sheets due the lack of depositional data in context sheets. Hence, these testimonies were classified as unreliable sources (Graphic 5.3). Overall, the analysis of depositional data demonstrated two things. First, it confirmed that investigating the formation processes of sedimentary layers can be more challenging than with masonry remains. Secondly, it demonstrated that a more complex recording system requires following more rules and standards. However, this doesn't imply that the quality of records will automatically improve specially if fieldworkers are not sufficiently skilled to achieve the task that the system demands.

As explained before, unreliable layer and structural claims can be considered as two additional forms of failures to interpret formation processes. And this information can be useful to make a general evaluation of the credibility of the sample in basic interpretations. This general index indicates that around 60% of the sample provides reliable sources, whereas 40% shows validity problems. However, these numbers obscure important facts, because around one half of the sample includes reliable context sheets (50%) and one third are unreliable sources (30%). On the contrary, one tenth of the sample represents reliable feature sheets (10%) and another tenth (10%) represents unreliable feature records.

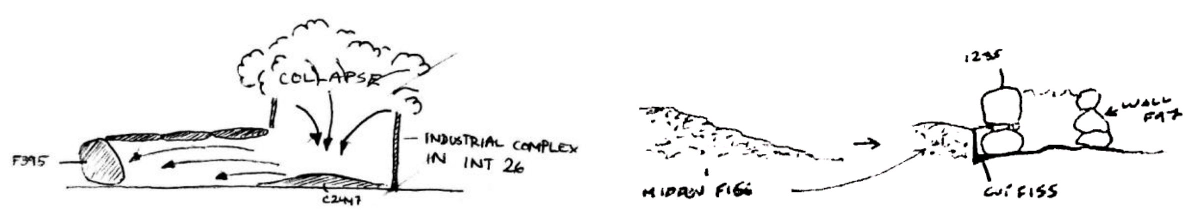


Figure 5.20. Sketch of the possible depositional process of layer C2447. (F395-LG) (left) and a sketch of the cut related to the constructive process of wall F97 (right). (F97-CKH).

Tarbat's sample also included some examples of evidential reflexivity when investigating formation processes (Graphic 5.4). Many of these examples describe doubts when interpreting the depositional process of a specific layers that could produce some ambiguity in the depositional history of features. For instance, the context sheet of a fill related to a flue asks: "*is* *this context evidence for the use of F395 or does it stem from a collapse of industrial feature?*" (C2447-LG). However, the main discussion was recorded in its feature sheet (Figure 5.20).

> *Backfill or sinkage from turf layer above? C2445. It is possible that C2447 forms part of the disuse of F395. Frog bones were found in C2445. Does this indicated the structural has been left open for some time after its disuse?* […] *C2447 is a charcoal filled deposit with an ashy appearance. This context may be associated with the collapse of industrial complex found in INT26, where ashy and charcoal filled material would be forced back out thorough the flue. It is also however possible that C2447 is associated with the use of F395" (F395-LG)*

Additionally, there was an atypical example of fluidity which describes an interpretive reconsideration in the constructive process of a feature. Specifically, this was classified an instance of evidential fluidity because the discovery of a cut led to the revision of a retrodictive conjecture about the constructive process of the wall (Figure 5.20).

> *"A drystone wall which appears to have no cut. Therefore, F97 could have been sunk/pushed into the soil to bed it, rather than a trench…. The western face of F97 -C1285- is contained/abuts the south edge of F155. This cut appears to represent the cutting of a shallow terrace into deposits below to provide a level base to the wall" (F97-CKH/ST).*

Finally, there was a small number of inadequate reflexive testimonies which only included short notes indicating ambiguity to interpret the depositional process of a layer: "*backfill/silting*" (C3541-NJT), but without providing any explanation or discussion. Some of these examples included an incomplete discussion providing some reasons for an interpretative possibility but failing to explain the other as in the following example: "*C1906 maybe a backfill that has washed away downhill, accumulation at linear F374 lowest point or silting up during disuse*" (C1906-LOB). Finally, it was observed that the best reflexive testimonies were often related to masonry remains whilst problematic testimonies were frequently related to sedimentary layers. This contrast was particularly evident in the records of one fieldworker that was quite successful to investigate the remains of a stone flue (Figure 5.21) but at the same time had much troubles to investigate a large pit with sedimentary layers (Figure 5.18). This represents one of the best examples that demonstrates that investigating a well-preserved masonry structure might be easier than a negative feature with more marginal evidence which requires more skill to be perceived and interpreted.

Y3.0	FEATURE RECORD	Site: TR01	Feature: 395
Intervention: 14	Quadrant: B7		

Easting: 883.64	Highest point: 14.09 m AOD
Northing: 998.71	Lowest point: 13.43 m AOD

Final form: Shape at top: SUB-LINEAR — Shape at bottom: CONCAVE

Profile: UNSEEN — Set with:

Description: F395 WAS FIRST IDENTIFIED AT HORIZON AS A LARGE SUB-LINEAR FEATURE CONSISTING OF TWO ROWS OF LARGE MIXED STONES (C2446) TERMINATING IN ONE LARGE ROUNDED STONE TO THE SW. THE FEATURE HAS ALSO BEEN CAPPED BY 8-10 LARGE STONE SLABS (C2359) UPON REMOVAL OF THESE STONES A VERY DARK SILTY SAND DEPOSIT WAS REVEALED. THIS WAS THOUGHT TO BE RECENT SINKAGE AND A CONTEXT NUMBER WAS THEREFORE NOT ALLOCATED UPON REMOVAL OF THIS LOOSE DEPOSIT A MORE CONSISTENT SOIL OF SIMILAR COLOUR WAS ENCOUNTERED (C2445) THIS HAS BEEN INTERPRETED AS A BACKFILL OF F395 (FROM TURF ABOVE?). C2445 OVERLAY A MOTTLED ASHY DEPOSIT WITH FREQUENT INCLUSIONS OF CHARCOAL, VERY DARK GREY IN COLOUR AND LOOSE IN COMPACTION. THE EXCAVATION OF THIS DEPOSIT DID NOT END UNTIL A COMPACTED, GREY SURFACE HAD BEEN REACHED (IT IS THUS POSSIBLE THAT THE FEATURE HAS BEEN OVER-CUT). F395 IS LINEAR IN SHAPE AND PTO

Stratigraphy	Evidence for
Overlies: Cuts: Abuts: Made of: C2446, C2359 Same as:	Construction: THE SIDES OF THIS FEATURE ARE MADE UP OF TWO ROWS OF LARGE STONES WHICH TERMINATE IN ONE LARGE STONE TO THE SW (C2446) THESE TWO ROWS HAVE THEN BEEN CAPPED BY A NUMBER OF LARGE, FLAT SLABS (C2359) THE DEFINITION IN THE AREA HAS NOT BEEN GOOD ENOUGH TO DETERMINE THE OUTLINE OF ANY PTO
Filled with: 2447	Use: C2447 MAY BE ASSOCIATED WITH THE USE OF F395.
Sealed by: Cut by: Abutted by: Backfilled with: C2445 Sinkage from:	Disuse: BACKFILL OR SINKAGE FROM TURF LAYER ABOVE? - C2445 IT IS POSSIBLE THAT C2447 FORMS PART OF THE DISUSE OF F395 (SEE COMMENT) FROG-BONES WERE FOUND IN C2445. DOES THIS INDICATE THAT THE STRUCTURE HAS BEEN LEFT OPEN FOR SOME TIME AFTER ITS DISUSE?

Comments: THIS FEATURE MAY FORM PART OF THE INDUSTRIAL COMPLEX CONTACTED IN INT 26. ALTHOUGH IT HAS BEEN SUGGESTED THAT F395 IS A FLUE OR AIR VENT, THE SPECIFIC PURPOSE OF THIS FEATURE IS AS YET UNKNOWN. IT IS POSSIBLE THAT THE DARK LAYER OVERLYING F395 (SEEN IN W FACING SECTION OF MOD B7) WAS PART OF THE FEATURE CREATING AN AIRTIGHT STRUCTURE SUCH AS A FLUE.
C2447 IS A CHARCOAL-FILLED DEPOSIT WITH AN ASHY APPEARANCE. THIS CONTEXT MAY BE ASSOCIATED WITH THE COLLAPSE OF THE INDUSTRIAL COMPLEX FOUND IN PTO

Date recorded: 27.08.01	Belongs to structure:
Recorded by: LG	Identified as: ? FLUE

Figure 5.21. Feature sheet describing a flue and its depositional history (F395-LG).

5.5.11 Groups of features, Structures, Spaces and Buildings

Groups of features, middle level contrasts, spatial and building data are commonly recorded in feature sheets (Graphic 5.2). However, the feature system uses groups of features as an auxiliary concept to define structures, spaces and buildings. As explained before, feature sheets include two sections for groups of features. One in the top of sheets labelled '*set with*' in the numbers of grouped features are listed (Figure 5.15). In general, these sheets also included an explanatory account to clarify whether a group of features defined a structure, a space or a building. For example, some sheets interpreted timber structures from the spatial arrangement in a group of stakeholes (Figure 5.22).

<u>Example 5.11</u>

> *"F561 to F571 were allocated to a series of 11 stake holes measuring between 0.06 and 0.15m in diameter and up to 0.15m in depth. These form a rough rectangular shape-surrounding hearth F535… Appear to relate to a structure over the early form of hearth F535 which may have formed a small furnace" (F561-NJT).*

Additionally, some groups of features defined internal areas according to the spatial arrangement among different types of features: "*A regular curvilinear gully forming a semi-circle around the southern and western sides of hearth F35 … The nature and position of this gully suggests its use as a wind break around the metalworking features such as hearth F35*" (F23-RTJ). This example is similar to the interpretation of rooms at B49, but Tarbat's example is based in more marginal evidence. Tarbat also included an example of an external activity areas interpreted from the recurrent appearance of features in one area (Figure 5.23).

<u>Example 5.12</u>

> *"F148 appears to form part of a metalworking complex two fired clay features (possible furnace bases F353, F479) lie to the south east and south west and a series of fired clay dumps (C2547) are located in close proximity. One dump abuts the stone of hearth makeup. All have produced suggestive evidence of bronze working" (F148-ST/NJT)*

Figure 5.22. A group of stake holes defining a timber structure around hearth F535 (Copyright Tarbat Discovery Programme). https://doi.org/10.5284/1031216

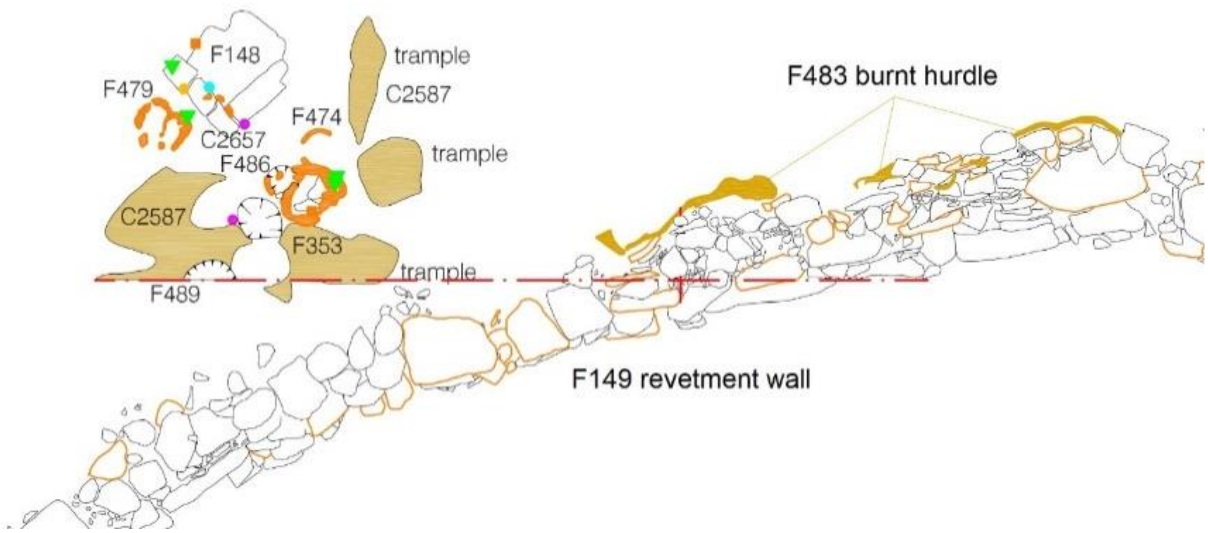


Figure 5.23. Metalworking area defined by the clustering of three hearth features F148, F353 and F479 (Copyright Tarbat Discovery Programme). https://doi.org/10.5284/1031216

Additionally, other feature sheets included information about groups of features in the section tagged as ‘belongs to structure’, this area is specially designed for linking features which define a building (Figure 5.2). In this case, the sample included various sheets of postholes, surfaces and a hearth that only included a structure number (S9) which corresponds to the vellum workshop

(Figure 5.24). However, none of these sheets included any explanation of the meaning or the evidence of this building. Instead, the interpretive process of S9 is better related in the annual reports (Spall 2006, 7, 2007, 10). This produced some doubt regarding the stage when this building was interpreted, but one of the supervisors confirmed that S9 was defined onsite. Hence, these examples demonstrated that defining a group of features onsite also requires explaining whether such group defines a structural element, a space or a building. In this way, the proportion of reliable group of features data indicates when this task was achieved, whereas the proportion of unreliable data indicates when this explanation was not included (Graphic 5.3). In general the proportion of invalid group of feature claims includes the sheets from S9 which deserved a detailed explanation, being an analogue case to Star Carr's house (Figure 1.6).

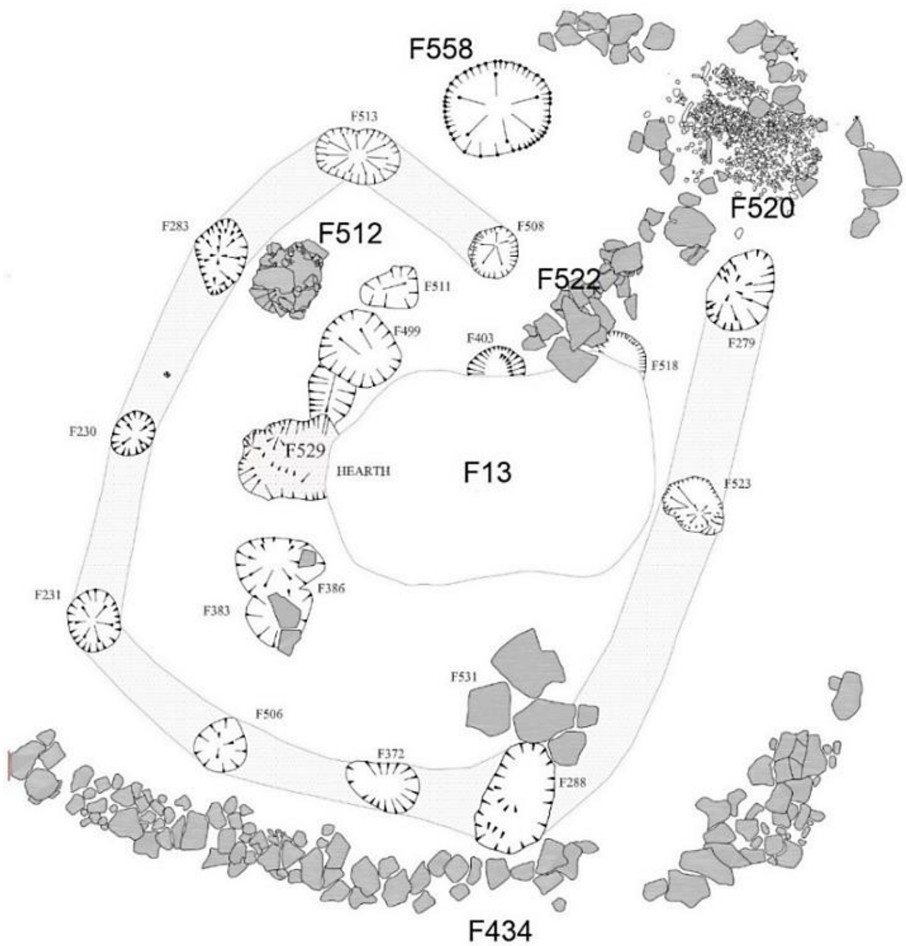


Figure 5.24. Structure (S9) or the vellum workshop. (Copyright Tarbat Discovery Programme).
https://doi.org/10.5284/1031216

Additionally, there are some sheets which report some form of spatial relation among features, but without defining a group or an area. For instance, one sheets describes the relation between a surface of cobbles and S9 which defines some sort of access to the building: "*Appears to lead from cobbled surfaces F520 down towards the interior of S9, may represent the stone paved threshold for S9*" (F522-RGP) (Figure 5.24). However, the sample also included some testimonies that describe a spatial relation among features but without enough explanation of their interpretive implications. For instance, one sheet mentions a relation between a posthole that "*may be directly associated with hearth 495/F529*" (RGP-532) but says no more. Hence, these cases were classified as instances of invalid spatial data (Graphic 5.3). There were some cases in which the credibility of groups of features and spatial data could not be achieved because only one of the associated features had been collected and to confirm the absence of an explanation required checking all the related sheets. But in general, only one of the related sheets included an explanation (Figure 5.15).

The sample provided a few examples of reflexive testimonies for spatial and groups of features. In general, most of these accounts describe reflexive situations when interpreting structures defined by a group of features. For example, the fragment below initially questions the spatial relation between an alignment of stakes (F404) and a wall (F394), therefore doubting whether these constitute a group of features.

> *Many suggestions have been made concerning these stakes; I feel that they are not structural in the sense of being associated or part of the nearby walls. If they were parallel with the banks of the channel, then you could suggest that they were there to revet. However, the fact that they are aligned with F394 suggests that they are concerned with what F394 does to the flow of water. Are they a fish or an eel trap perpendicular to the water flow? (F404-TJS)*

Additionally, there is one example that relates various interpretative reconsideration when investigating a group of features (Figure 5.25). Initially, F436 was described as a fence line interpreted from a small trench with remains of wooden posts except for a small gap in the middle, but later the digger observed that this gap was "*echoed by the gap in two features to the south, F442 and F443 which lie parallel to F436… it appears that F436 is contemporary in function and time as southern gullies F442 and F443*" (F436-CKH). Then, it was considered that "*F442 and F443 could be a possible fence gully line of some kind… although no wood or* *stakeholders were found*" (F442-CKH). Finally, after excavating F443, the digger considered an additional piece of spatial information to reformulate the interpretation of evidence as a group of parallel fences defining a gate access.

> *"F443 runs slightly off parallel to F442 a similar sublinear gully but ending approximately 0.4m from the terminal of F443. These gullies form a probable contemporary function with the gap some sort of access or gateway. This gap is echoed in F436 the fence gully to the north. Within this gully are two trenches containing vertical wooden posts with a gap between at exactly the same place as the butt ends of F442 and F44. Although it is unlikely that F442 and F443 would have been utilised as water management, to the south an ancient stream lives only 3m away, therefore suggesting these features acted as a barrier (stock control) from the stream edge" (F443-CKH)*

This example is interesting because the sequence of sheets describes the formulation of an explanation for that group of features. Nevertheless, the interpretation is always uncertain due the fragmentary character of remains. As one can see most of these doubts are produced by marginal or fragmentary evidence that produces uncertainty when interpreting the character of remains, therefore representing cases of evidential reflexivity (Graphic 5.4). Finally, one also must notice that Tarbat showed a wide diversity when defining groups of features, spaces and buildings and that many of these data have been interpreted from marginal or fragmentary evidence. Nevertheless, one should not ignore how the excavation strategy, the pace of excavation and the character of the deposit also contributed for the richness of these data (Figure 5.5).

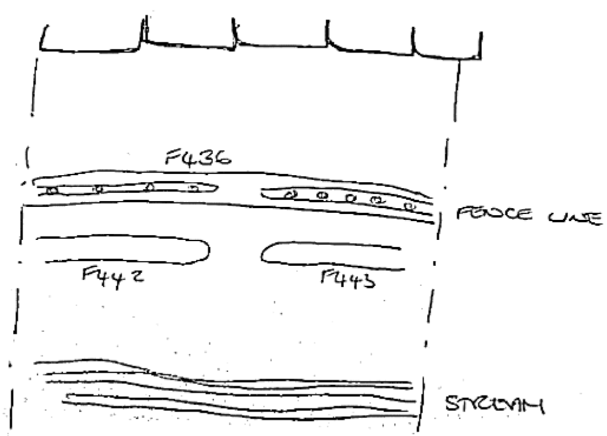


Figure 5.25. Gateway access defined by F436, F442 and F443. (F443-CKH).

5.5.12 Phases

Tarbat records only included a few instances of phase information (Graphic 5.2). However, all these claims are indirect references like in the following fragment: "*although this pit belongs to the phase of metalworking, there is no evidence to suggest it was associated with this activity*" (F13-AR). In other words, none of these testimonies provides an actual description of phases and therefore no further evaluation was executed. Complementarily, a fast revision of excavation diaries[9] a few detailed accounts of site phases as the example below which describes one of the latest phases of the site. This entry was accompanied with a sketch that facilitated the interpretation of the written information (Figure 5.26). This document confirmed that phase investigation had been better developed onsite and provided additional evidence that notebooks can be useful tools for recording this type of synthetic claims. However, a systematic analysis of site notebooks was impossible due the lack of a systematic diary record which only covered the first three seasons. This also demonstrates that having an additional tool is not enough for having higher quality testimonies if elementary recording rules are not followed. Additionally there is published testimony from the project director that resumes phase definition onsite (Carver 2016, 68).

> *"The latest phase in quadrant D is characterized by four concentrations of dump shell middens. One is at the centre of the north end of the quadrant and on the edge with INT14. The second is at the centre of the quadrant, the third is next to it and continue under the eastern section, the last one is on the south west corner. A drystone wall (F97) runs south north on the eastern part of the quadrant and seems to have a turn of the corner towards the east (F48). This wall appears to be earlier than shell middens as shells are located on top of it. Site notebooks Sector 2. 31/07/97*

[9] Excavation diaries are catalogued as 551 359/76. Unit ID 33106

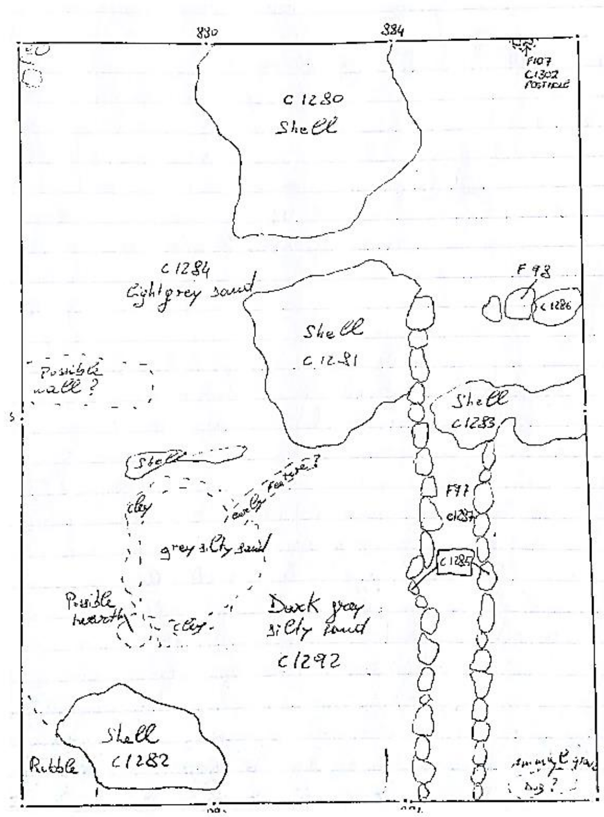


Figure 5.26. Preliminary sketch from an excavation diary representing an occupation phase at TRS2 defined.

5.5.13 High Level claims

High-level claims include diverse types of interpretations: functional, agency and culture/period. Although, these are not very frequent (Graphic 5.2). The sample included a small number of agency claims which are very similar to previous case studies, for example dump events reinterpreted as cleaning action. However, one of the most distinctive examples describes the interpretation of various dumping events in a shell midden as evidence of seasonal activity (Example 5.10). In this way, depositional evidence is reinterpreted in terms of human action:

> *"F156 allocated to a series of dumped shell contexts …comprises a layer of mainly intact winkle shells C1348 overlying a more extensive, largely smashed layer of mussel shells C1427, itself overlying a second, less extensive winkle spread C1428. Animal and fish bones were recovered from all context […] Layering effect of different shell species suggest perhaps a seasonal chance in species caught…Soil accumulations above is slight, inferring a period of inactivity of shell dumping (possibly seasonal)" (F156-RBM)*

Culture/period statements are not very frequent and generally they show similar recording pattern previous case studies. For example, when describing the fill of a drain, one sheet mentions: "*contained a discrete dump of smashed Pictish sculpture*" (F180-CKH/CAS). Likewise, when describing a group pf features, a sheet mentions: "*these features are possibly pre 8th century"* (F443-CKH) (Figure 5.25). The low frequency of these claims seems to be consequence of the character of the site that was mostly restrained to the early medieval period (Carver 2016, 46).

As in previous case studies, functional data is the most frequent high-level information and it is commonly recorded in features sheets which often describe the function of features and structures. In general, there is a good balance between structural and non-structural features. Many of these examples are interesting because although remains are not fragmentary, evidence is marginal. For example, the interpretation of a storage pit is partially based in the location of the feature with S9: "*This pit is immediately north of a possible entrance to structure (the leather workshop where winkles are being burned in a hearth. It may have been a supply store for this purpose*" (RS-F558-RS). But additionally, the fieldworker noticed crucial evidence when investigating a fill of shells and observes that:

> *"winkles in brown sandy matrix, at the top the shells were broken and decayed but below this were intact and well preserved…possibly this was a store of shells ready for use in structure, where we know they were being burnt" (C3491-RS).*

The sample also includes more typical examples where fieldworkers interpret the function of features considering find's evidence like in the following example: "*slag and waste of bronze products in the backfill and the lining to the pit suggest its use as a bronze working pit*" (AR-43). Likewise, there are some examples where fieldworkers consider marginal evidence in nicely preserved masonry features like in the following example for a stone-lined pit.

> *"A water collector feature for some purpose…the feature is at the lowest part of the site, the base would have been quite water bright, particularly if made up of one or few flags and the rather irregular lining could have been semi-water tight if turf was interjected with the stones" (F470-RS).*

In general, functional and agency claims are justified quite often (Graphic 5.3). This indicates that when diggers reached this stage of interpretation, in general site records are carefully recorded; although, this stage was not reached very often. As in other cases studies, culture/period statements commonly were unjustified, but these were not counted as invalid testimonies given the methodological characteristics of the project divided among excavation and post-excavation work. Finally, there are few reflexive claims, but these are always related to functional aspects, more commonly for non-structural features. The example below represents one of the best examples of reflexive discussion regarding the use of a large gully ditch, besides this is one of the fewest examples where cut's evidence is discussed.

<u>Example 5.13</u>

> *"First possibly thought to be for drainage however backfill or base revealed no wash/waterborne silting and therefore this is not possible. MOHC suggested that this gully and the two adjacent gullies F172 and F59 were probably used as a cattle or people drove after the cobbled road (F18) was in disuse. However (ha ha) this idea does not fit with the 'perfection of the cut', the sides cut rather than trampled. Interestingly F180 and the two adjacent gullies do not appear to have been in use at the same time, possibly suggesting that they served similar purposes and for the same reason no longer suited this use and therefore where replaced by another, being backfilled deliberately with soil and large stones" (F180-CKH/CAS).*

Additionally, there were some cases of reflexive data which describe functional uncertainty, but they often lacked a discussion of evidence, therefore being classified as instances of problematic reflexive data: "*Uncertain function. Possibly a surface/stance associated with leatherworking*" (F525-NJT/DMK) or "*probable water collection/rubbish pit*" (F528-CAS) (Graphic 5.4). Finally, there are a couple of sheets in which fieldworkers claimed a lack of functional evidence like in the example below. However, this testimony was not very credible because the recorded had failed to interpret the type and formation processes of the feature (Figure 5.18). Still, this testimony was still useful to infer that his claim was most likely the result of limited interpretive skills for the investigation of F345.

<u>Example 5.14</u>

> *"The shape and profile of F345 are unusual and distinctive. No evidence for function however could be identified in section plan or backfill. No evidence for any structural components were found likewise found" (LG/ST-F345).*

Moreover, despite the problems mentioned above, this record included relevant descriptive data that allow to reinterpret the feature offsite. For example, the description of the cut mentions:

> *"F345 had a distinctive u-shaped profile with sides sloping along its length at between 70-85° bottoming out into a slightly concave base. At the terminal ends of this elongated pit, the slope was more gradual 35-60°" (LG/ST-F345).*

Whilst the description of the main backfill described

> *"A very dark brown silty sand deposit located in the SE corner if module B0. Inclusions of flecks of charcoal, burnt and unburnt animal bone, unburnt fishbone, burnt and unburnt sandstone and quartzite pebbles and slag. The deposit measures 2.5m length, 0.30cm depth and 40-45cm width" (C1827-LG)*

Altogether, this evidence suggests a possible rubbish pit considering the irregular shape of the cut and the variability of burned and unburned finds in the main fill that indicate dumping activity (Figure 5.27). In this way, although this testimony was not very reliable in multiple aspects, its accurate observations allowed the reinterpretation of the document. Overall, this was a perfect example that the value of a documental analysis.

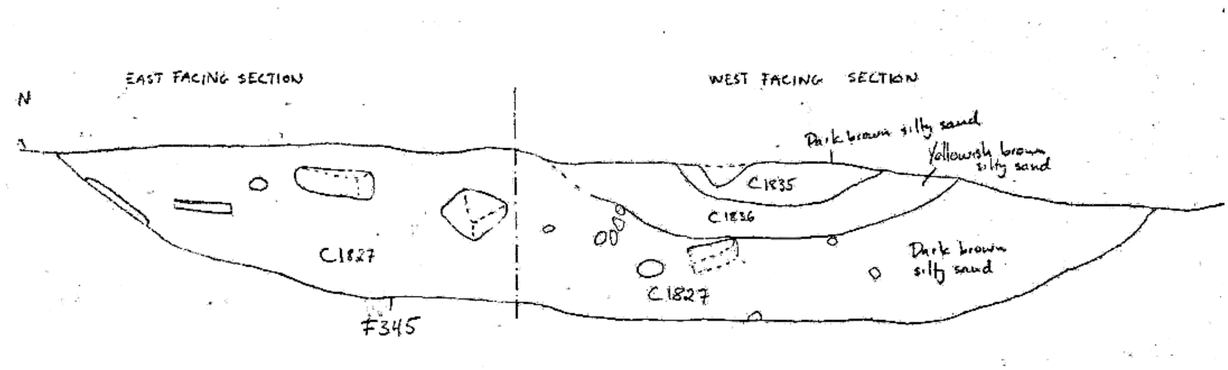


Figure 5.27. Section record for F345 (LG-F345).

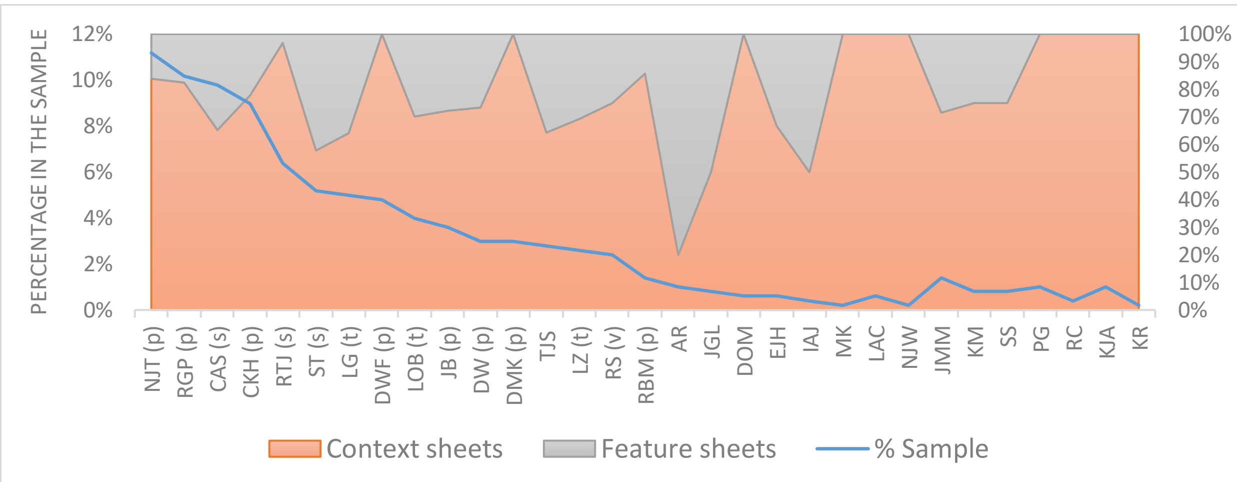
5.6 Digger's performance

Tarbat had a mixed workforce that included professional diggers, trainees and volunteers with different ranges of participation in the project. Table 5.1 describes the participation of sixteen fieldworkers considering the number of worked seasons. This group of fieldworkers was defined as the 'main team' and the records of these participants represented 84% of the sampled. The proportion of sheets recorded by each participant of the main team varies between 2% and 12% (Graphic 5.5). The rest of the sample includes records from fifteen fieldworkers and the proportion of sheets for each one is very low (0.6% average). In general, this includes records by trainees and volunteers that only participated in one season, including a few sheets recorded by two supervisors with a more restricted participation in recording (AR and JGL). These have been defined as the 'backup team' and the records of these participants cover 16% of the sample.

Digger	TR96	TR97	TR98	TR99	TR00	TR01	TR03	TR04	TR05	TR06	TR07
NJT (p)											
RGP (p)											
CAS (s)											
CKH (p)											
RTJ (s)											
ST (s)											
LG (t)											
DWF (p)											
LOB (t)											
JB (p)											
DW (p)											
DMK (p)											
TJS (p)											
LZ (t)											
RS (v)											
RBM (p)											

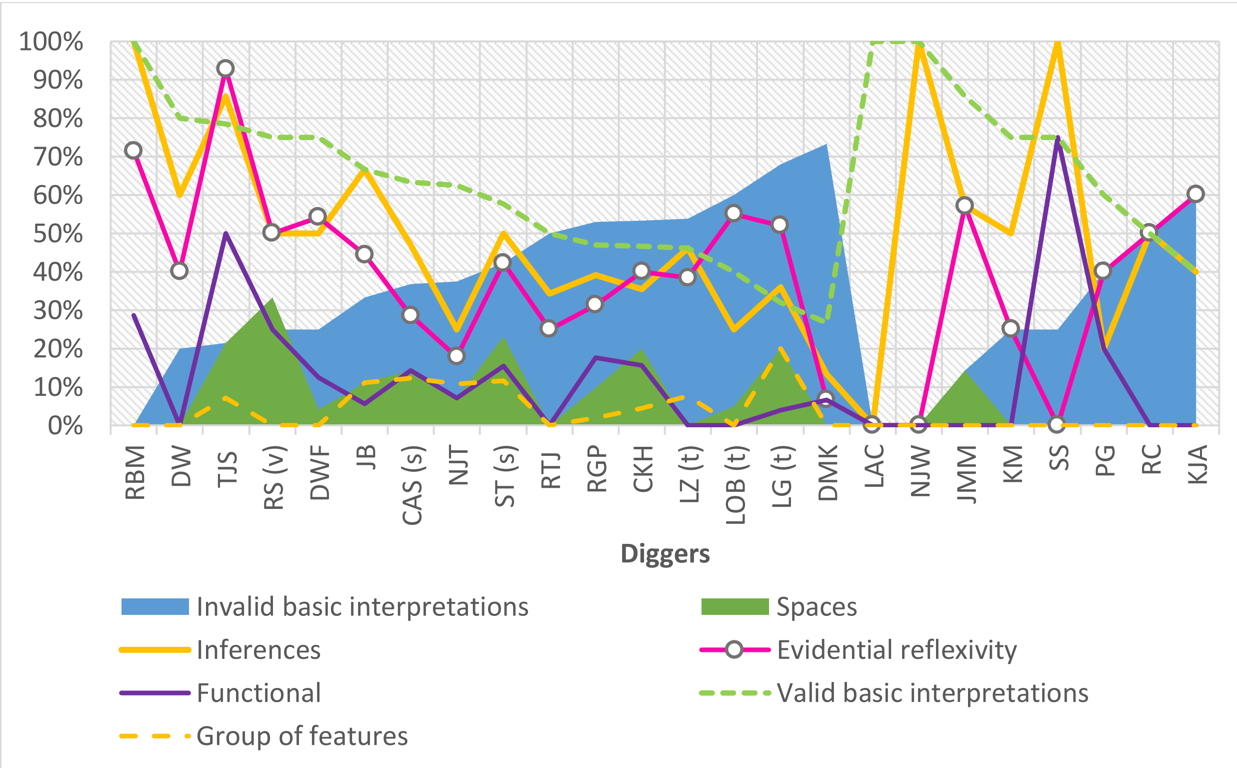
Table 5.1. Participation in the project of the 'main team'. Tarbat. The main team included the participation of professional diggers (p), trainees (t) and a volunteer (v) with an extensive participation in the project. Highlighted with similar colour are diggers that joined approximately at the same time.

Graphic 5.5 also shows the proportion of recorded context and feature sheets by each participant of the main and the backup team. Interestingly, this data shows that some fieldworkers didn't participate in feature description, for example two diggers from the main team (DWF and DMK). Although, it was more common that fieldworkers from the 'backup team' had less chances to participate in feature description (DOM, MK, LAC, NJW, PG, RC and KJA). This recording behaviour is consistent with the general excavation strategy because the exploration of a feature could last more than one season whereas many people in the backup team only participated for one season. However, professional fieldworkers had more access to recording work in general. In fact, professional diggers were commonly referees as recorders (Carver 2016, xvii). On the contrary, trainees and volunteers had more restrained access to recording work in general and this indicated a degree of separation between digging and recording activities. This was confirmed with some testimonies like the following example in which a professional fieldworker describes the work made by a student: "*F166 was excavated out of sequence at its western limit to recover sculpture 1798. Once the stone had been recovered and removed this remaining sample was excavated by second year student APW*" (F166-ST). Nevertheless, there might be good reasons for restraining trainees and volunteers from feature description which after all is a more complex. However, some testimonies also showed a more positive was approach when professionals checked, and corrected feature sheets filled by trainees (Figure 5.19).



Graphic 5.5. Distribution of the sample by fieldworker. The graphic also indicates the distribution of context and feature sheets by participant. Main team {RBM…LG} and Backup team {AR...KJA}. Professional diggers (p), trainees (t) and volunteers (v)

As in previous case studies, the main goal of the individual analysis was to examine the differences and similarities in the interpretative practice of fieldworkers. Graphic 5.6 illustrates the percentage of reliable and unreliable basic information for each participant of the main team (RBM…DMK) and some participants of the backup team (LAC…KJA). However, the trend for valid data partially correlates with the trend of Inferences and Evidential reflexivity. In other words, the less interpretive mistakes a digger makes, the more able he's to make interpretive conjectures and to be record relevant reflexive data. This graphic clearly shows that trainees in the main team (LZ, LOB, LG) produce unreliable data more frequently. Results for the backup team (LAC…KJA) show a similar distribution between valid and invalid claims, but the proportion of inferences and evidential reflexivity is more irregular, which indicated a much less systematic field practice. Finally, another important difference is a more recurrent presence group of features, spatial and functional data in the main team. In sum, this graphic describes a better performance among trained fieldworkers and those with a longer participation in the project. In contrasts, the less skilled fieldworkers and those with a shorter participation had less chances to achieve a systematic practice.

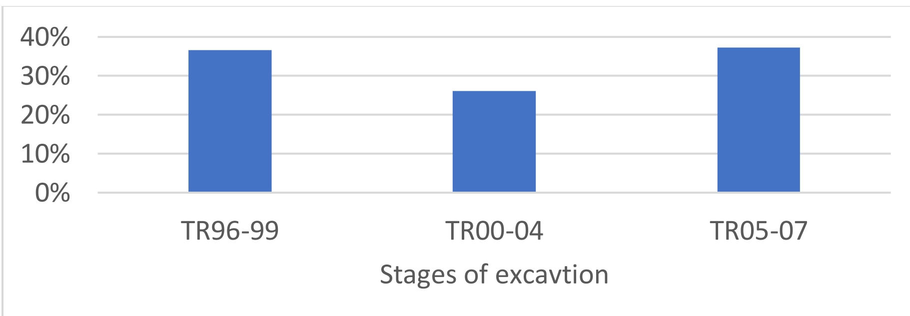


Graphic 5.6. Diggers' performance. Tarbat. The graphic illustrates the distribution of valid and invalid basic interpretations for the participants of the main team (RBM…DMK) and some participant of the 'backup team' (LAC-KJA). The graphic also includes the distribution of Inferences, Evidential reflexivity, Space, Group of features and Functional claims.

Tarbat records also provided sound evidence of collaborative interpretation. One of the most interesting examples was the interpretation of a metalworking area which began with the definition of a hearth F148 in 1999 described by ST (Example 5.3). However, when the exploration of the area continued in 2004, NJT found two more hearths (F353 and F479) and defined a metalworking area considering the recurrent type of evidence (Example 5.12). Even though, ST and NJT didn't work together (Table 5.1), the interpretation of the metalworking was derived from the observations made by ST and NJT (Figure 5.15).

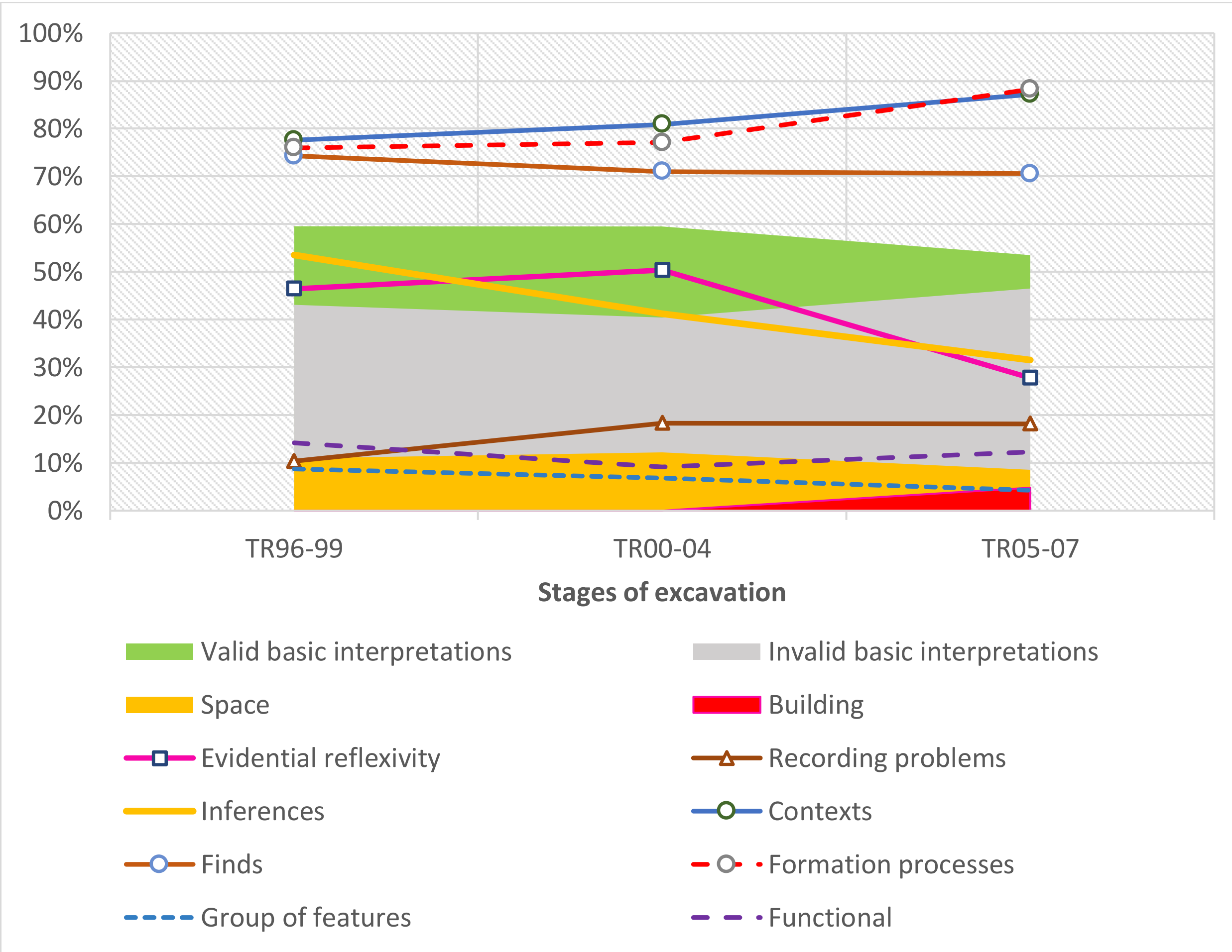
5.7 Longitudinal analysis

Tarbat's exploration lasted twelfth seasons between 1996 and 2007, with no activity in 2002. Each season had an approximate duration of two months, thereby the sample covers around 24 month of work which represent the longest examined excavation. To keep consistency with previous case studies, the sample was divided in three phases to examine the distribution of claims (Graphic 5.7). The initial phase covers the first four seasons of the project (TR96-TR99), the middle phase covers the following four (TR00-TR04) and the final phase covers the last three seasons (TR05-TR07).



Graphic 5.7. Distribution of the sample into three arbitrary stages. TRS2.

Graphic 5.8 shows the distribution of claims cross the three phases. In the top of the graphic, one can observe a regular distribution of information most systematically recorded (context, finds and formation processes). Down to the bottom, the trends for spatial, group of features and functional data also show a consistent appearance, although the frequency is much lower. The consistency of spatial data and group of features are specially interesting because these seem to confirm the positive effects of the excavation strategy. Yet, building data is one exception because this information only appears in the final phase of the project, therefore indicating that this type of claim requires an accumulation of background knowledge produced in the previous phases. Finally, the graphic indicates the frequency of valid and invalid basic data along the history of the project. This distribution shows a similar proportion for reliable and unreliable testimonies in the first two phases. However, there is a slight reduction in the quality of records in the last stage which is correlated with a decrement in the number of inferences and evidential reflexivity. Overall, this graphic demonstrates a systematic practice in the capture of various types of data. However, some more complex interpretations like the definition of S9 depend on time which enables more familiarity with the site. But at the same time, there is a decrement in the quality of records possibly related to the increasing pressure as the project comes to an end, although the multiple changes of staff during the period between TR03-TR05 (Table 5.1).



Graphic 5.8. Longitudinal distribution. Tarbat. This graphic illustrates the distribution of valid and invalid basic interpretations, and various types of claims over the course of three arbitrary stages of the project.

5.8 General observations

<u>Recording strategy</u>. The feature system is an alternative method that aims to capture more information than SCR. However, this is not very clear because the information recorded in both systems is very similar. In that sense, the feature system should be better described as a different way to report data rather than completely different ways to approach interpretative work. Still the feature system has some methodological advantages because the feature sheet facilitates the

organization of analytical data in context sheets and synthetic information in feature sheets which facilitates the analysis of records overall. Another important virtue of the system was a better strategy to report spatial data based in a systematic use of explanatory accounts. At the same time, the method revealed various problems as the lack of a cut's sheet, a finds section in context sheet and some important ambiguities in stratigraphic sections. Additionally, the procedure for defining and describing features was not very well stablished.

Tarbat showed a lower number of sketches, however these achieved similar functions as those observed in London. In some cases, sketches contribute to illustrate operative decisions (Fig 5.6) and reflexive situations (Fig 5.16). And in other cases, sketches provide visual evidence for interpretive claims (Fig. 5.14). In general, there was a more limited use of sketches to illustrate spatial data and groups of features. For example, no sketches in primary records of the interpretation of S9 and the metalworking area were found, which would have worked very well to complement explanatory accounts. A distinctive aspect of Tarbat's sample was the use of sketches to illustrate formation processes (Fig. 5.19).

<u>The interpretative process</u>. As mentioned before, in general terms the distribution of interpretative claims is very similar to former projects. Yet, one of the characteristics of Tarbat was its broader variability of spatial data. However, this success should be understood in relation to the excavation strategy and the character of the site. Additionally, the analysis in the reliability of data demonstrated that having a more complex system will not automatically improve the quality of records because a more complex strategy demands following more rules, hence these creates more spaces to mistakes especially for untrained fieldworkers. Finally, the internal analysis also proved the existence of an important epistemic change in the excavation phase of the project as the team achieved a more detailed functional picture of the site in relation to its evaluation stage. Particularly, in the industrial activities during the monastic occupation.

<u>Reflexivity</u>. In general, Tarbat records report reflexive data following the same strategies observed in previous case studies. Yet, Tarbat pay more attention to evidential aspects, although a recurrent problem is that relevant data are not properly reported very often.

<u>Non-interpretative information</u>. In addition to the more systematic and repetitive non-interpretative content, there is a smaller amount of site records that include additional circumstantial and operative information to explain specific procedures in the exploration of a

specific layer or feature. Overall, this additional information had an important function to explain if these procedures affected or not the credibility of data. Therefore, achieving a similar function than reflexive diaries. Once more, it was confirmed that despite the limited amount of sampling information, these decisions can be properly understood when interrogated against descriptive and interpretive information.

Individual and longitudinal analyses. The analysis of the individual performance confirmed the incidence of skill in the reliability of data. However, the individual analysis also confirmed that participation is another factor that contributes for the interpretation of higher-level data as spatial information. Besides, the analysis of individual performance also demonstrated a restriction of recording activities for less experienced participants. In that area, Tarbat was a les democratic project, yet when trainees had access to feature description, they produced unreliable records more often. Then, as explained before a good balance between a hierarchical approach should give more recording opportunities to trainees but they should be guided by a more experienced participant, especially to minimize errors. The longitudinal analysis showed a very regular practice in many areas of investigation including spatial aspects, except for the building interpretation of S9 which also depended in the accumulation of background knowledge. However, with the course of time, there was also a negative effect in the credibility of data. Altogether, the individual and the longitudinal analysis provided different types of evidence that indicate that the production of good quality primary data is more complex than simply adding another sequence of sheets.

Chapter 6 . Discussion

---

field archaeologists dig up rubbish, theoretical archaeologists write it down.
Paul Bahn, 1989, 15.

Previous chapters have presented an individual analysis of projects. Then, this chapter summarizes quantitative and qualitative observations of case studies to stablish a comparative judgment of recording methods. Initially, there is a quantitative comparation of the distribution of data, but these results are interpreted considering the epistemic context of each project (6.1). Afterwards, there is a comparation of reflexive content in case studies, this examination is based in the quantity and relevancy of reflexive information (6.2). This is followed by a contrast of non-interpretative information based in the amount, the relevancy and the redundancy of this type of information (6.3). The comparative process concludes with a contrast in the reliability of records from SCR and the feature system. This section explores more deeply the most recurrent mistakes in testimonies and makes a final consideration on the importance of refining our ideas about documental analysis (6.4). The following sections resume the main virtues and problems in the formal aspects and guidelines for each method, hence this defines more clearly aspects that could be imitated and aspects that deserve further development (6.5).

Up to this point, the investigation has focused in describing the variability of samples but without paying much attention to the use of textual and visual testimonies. Hence, this chapter takes the opportunity to reflect on this aspect more extensively (6.6). Finally, a goal of this investigation has been to revise some aspects of recent interpretative theory, especially its strong emphasis in embodied and subjective performance. Hence, the next section summarizes the evidence that demonstrates the existence of an individual and collective dimension of field interpretation. Particularly, this section advances a more detailed characterization of collective interpretation (6.7). Finally, the chapter concludes with a brief description of the main methodological limitation of this investigation and advances some future avenues of research that could be taken on from the results of this project (6.8).

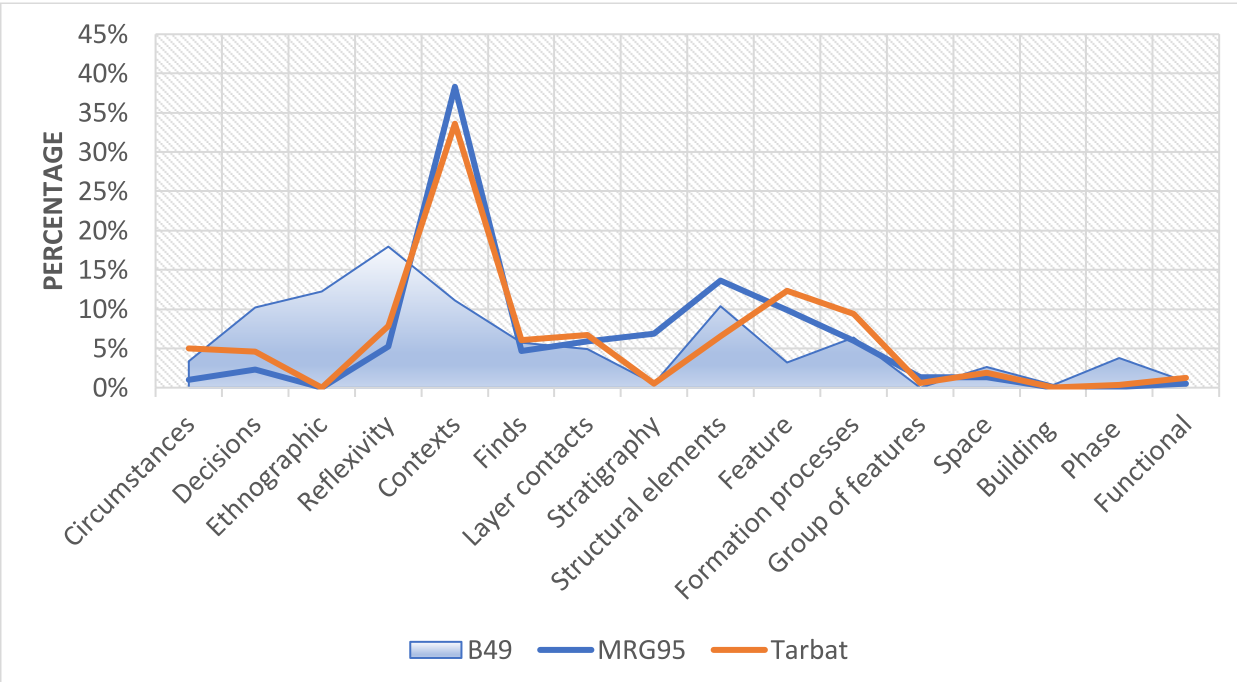
6.1 Comparation by type and epistemic context

The internal analysis allowed to map the distribution of claims in case studies and to observe the incidence of contextual factors in interpretative work, specifically the skill of fieldworkers, the character of remains and the background knowledge of sites. Hence, the purpose of this section is developing a comparative judgment of case studies considering these elements. A primary goal of this contrast is to examine whether one can observe more reflexivity and interpretation when using a more complex recording strategy (Hodder 2000c, 7; Farid 2000, 25; Carver 2009, 12). Specifically, this comparative judgment is based in a contrast of qualitative results of case studies, but always considering the epistemic context of projects.

Graphic 6.1 illustrates the distribution of the main types of information in case studies, particularly the variability and percentage of claims. In terms of variability, the graphic shows a similar collection of claims in every project, hence this indicates no important differences between methods and case studies as all of them report similar information. More specifically, there are two types of claims which are exclusively found in B49 records, namely ethnographic and phase claims. As explained before, ethnographic information is largely irrelevant and therefore this doesn't represent an important contribution to improve the quality of records. On the contrary, phase claims are more interesting because this is a higher-order claim. Hence, this is an aspect in which dairy records are superior Yet, to assume that this is exclusively explained by the recording format would be misleading. As observed before, phase claims are exclusively reported by professional diggers (Graphic 3.5). But more importantly, the character of remains also facilitated this type of interpretations, especially the repetitive constructive pattern in which a new phase was built on top of a previous one, generally following the original architectural plan. For this reason, constructive phases were more difficult to interpret at MRG95 because these changes were less evident. Finally, Tarbat results show some similarities with B49, as both illustrate the benefits of diary recording for phase claims, but the lack of systematicity constrains further comparations.

Group of features is another important element because it was absent from B49 (although this is not very clear in the graphic, given the small percentage of these values). On the contrary, MRG95 and Tarbat fieldworkers recorded different types of observations and conjectures based in the alignment and proximity of two or more features to define structures and spaces.

Comparatively, B49 diggers never had to consider this type of claims as most structures were largely evident and the spatial organization of the building was relatively well defined by internal walls. Thereby, none of these examples provides sound evidence of one project being more interpretatively effective than another. More exactly, evidence indicates that minor variations in the variability of data can be explained by differences in the epistemic context of projects.



Graphic 6.1. Comparation by type and case study. The graphic illustrates the distribution of the most recurrent types of claim in projects. Reflexivity summarizes the percentage of uncertainty, fluidity and possibility claims.

In terms of percentage, the graphic reveals greater similarity in the recording activity of MRG95 and Tarbat, but the most important element is an outstanding proportion of context descriptions. On the contrary, B49 shows a lower proportion of layer descriptions and a higher amount of reflexivity and decision claims. Thereby, this data seems to confirm that reflexive methods pay more attention to non-interpretive and reflexive information than standardized systems (Hodder 2000a, 9; Mickel 2015, 303). However, this assessment must be made considering that quantitative variations also obey to formal differences, particularly the extensive gloss of diaries (see section 3.4). And more importantly, one must consider the proportion of relevant and trivial of reflexive and non-interpretative data to make a more accurate comparative judgment (see Sections 6.3 and 6.4).

The proportion of structural, feature, formation processes, spaces, building and functional claims show great coincidence in the three projects. Nevertheless, this obscures important qualitative aspects. For example, MRG95 and Tarbat showed a higher diversity in the type of feature and formation process claims, whereas these aspects were highly uniform at B49 given the predominance of masonry remains at Catalhöyük. Moreover, B49 diggers rarely faced important interpretative challenges due the fragmentariness of remains whereas this kind of situation was more recurrent at MRG95 and Tarbat. Similarly, Tarbat showed a larger diversity of spatial claims, for example the definition of external areas, a building (S9), and spatial relation among features like the access to S9. Specifically, this is an aspect in which the dimension of the excavation area and the pace of exploration were very rewarding. MRG95, included similar examples as the definition of a nuclear activity area defined by the kilns and storing pits, an access to the site and some spatial relation between the site and the river. And equally important is noticing that most of these claims are based in marginal evidence. On the contrary, the presence of B49 was evident since the earliest stages of excavation and the definition of rooms was clearly indicated by internal walls too. Thereby, if one considers the diversity of interpretative situations and the fragmentariness of remains in each project, one can say with great confidence that the interpretative process at MRG95 and Tarbat was more complex.

Similarly, functional data indicate that MRG95 and Tarbat were more successful if one considers the epistemic change before and after explorations. In both cases, evaluation work indicated the industrial character of MRG95 (Museum of London 2000, 15–16) and Tarbat (Carver 2016, 27). Yet, excavation work contributed to develop a more refined pictures of the diversity of industrial activities. Particularly, site records show how this goal is slowly achieved by the investigation of features and structures, even though functional interpretation is very occasional. In contrast, B49 is not very successful demonstrate an important epistemic change because the team was unable to judge whether new evidence supported or contradicted the assumption about the domestic character of the building. One particular goal of the reflexive method was that fieldworkers could evaluate their preconceptions of the site and perhaps the Neolithic (Lucas 2014, 4019). Hence, this was a missed opportunity because most of this exploration only confirmed what was already known from surface and previous explorations, that is the characteristics of a building.

In sum, comparative analysis has not provided relevant evidence to subscribe that having a more complex recording strategy increases the variability and percentage of interpretative data. On the contrary, comparative judgment indicates that MRG95 and Tarbat were more successful investigations if one considers qualitative aspects like the complexity of interpretative tasks faced in those sites and the epistemic change before and after the explorations. Evidently, B49 demonstrated to be more successful for the interpretation of constructive phases, but this is explained by this is also explained by contextual factor and not by methodological aspects. It remains to be proved whether reflexive diaries provide more relevant reflexive and non-interpretative data.

6.2 Reflexivity in excavation records

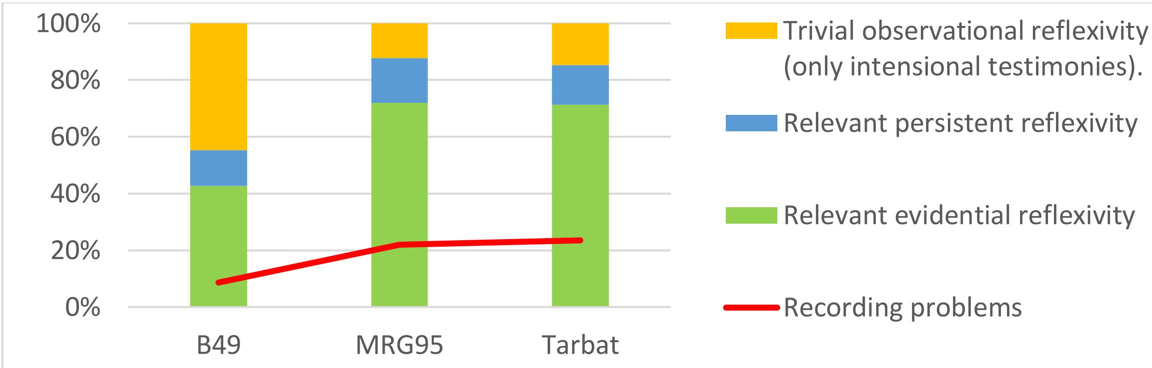
Another type of epistemic situation described in site records are reflexive scenarios which give account of doubts and reconsiderations in the process of interpretation. As mentioned before, a recurrent argument of reflexive archaeologists is that standardized formats fail to capture these aspects (Hodder 1999, 31; Mickel 2015, 303). Evidently, evidence demonstrates that such claims is false but more importantly, the study of reflexive information allowed to identify two classes of reflexive testimony. Observational and evidential reflexive claims. Cases of observational reflexivity describe interpretative doubts and reconsiderations in the process of discovery of remains. In general, these testimonies take the form of a narrative that describes an initial doubt regarding the type of vestige being defined, sometimes this is followed by episodes of interpretive reconsideration as clearance continues, but eventually uncertainty disappears when vestiges become visible. These testimonies are frequently found in reflexive diaries (Example 3.8 and Example 3.14) and less frequently in feature sheets (Example 5.5). However, these testimonies have been considered trivial or unnecessary for the credibility of primary observations because a description of vestiges, a sketch or a photograph can be a sound proof of having observed such things, especially if remains are well-preserved (Figure 3.3).

The analysis of context sheets from MRG95 identified some testimonies of observational reflexivity, but unlike previous cases, these were not textual. Specifically, some context sheets include sketches recording doubts and reconsideration related to the process of discovery (Figure 4.18 and Figure 4.20). Additionally, during the analysis of context sheets corrections in the matrix were interpreted as unintentional evidence of interpretive reconsideration in the process of discovery (Figure 4.15). An important characteristic of sketches and corrections providing reflexive testimony is having a heuristic function during the definition and correction of primary data. Hence, even if these testimonies are not fully useful in the phase of documental analysis, it is more difficult to define them as trivial. Besides, this evidence proves that the act of discovery is not totally expelled from standardized formats, but instead of using textual narratives, context sheets use graphic elements.

Additionally, the analysis of reflexive claims demonstrated the existence testimonies of evidential reflexivity. These examples describe interpretative doubts when evidence was puzzling for being marginal or fragmentary and therefore the credibility of an interpretation was compromised. Many of these doubts emerged when the type of structure (Example 3.9 and Example 5.4) or its existence was unclear (Example 4.12). Other examples describe cases when the existence of a layer (Example 3.3) or a feature was not conclusive (Example 4.9 and Example 5.9). Occasionally, records describe situations when it's unclear whether a find should be grouped into an assemblage (Example 3.12) or if a layer should be grouped into a feature (Example 5.10). Cases of evidential reflexivity also include examples of interpretive reconsiderations or fluidity when a retrodictive conjecture is revised due the discovery of new evidence (Example 4.11 and Example 5.13). Unlike observational fluidity that describes the process of belief formation, evidential fluidity represents cases of belief revision. In general, examples of evidential fluidity are not very frequent except at B49 and these constitute the main core of evidential reflexivity. In such cases, diggers usually revise previous ideas about depositional processes after receiving feedback from laboratories (Graphic 3.3). Particularly, this is one of the aspects where one can observe one of the strongest elements of the reflexive method (Example 3.13). Overall, evidential testimonies are considered relevant because they describe situations when ambiguous evidence affects the credibility of interpretations.

The analysis of evidential testimonies also revealed that the best format for these reports is an argumentative discussion of evidence and interpretative possibilities. Thus, when reflexive diaries include evidential testimonies, they change from a narrative to an argumentative style. Furthermore, evidential testimonies in context and feature sheets occasionally use sketches to complement textual testimonies, for example to illustrate ambiguities about features existence (Figure 4.25 and Figure 5.16) and grouping decisions (Figure 4.26). Nevertheless, a recurrent problem is that such recording standard is not always achieved when evidential testimonies lack an argumentative discussion. Finally, there are testimonies of persistent reflexivity which describe uncertainties and ambiguities produced by puzzling evidence but without compromising the credibility of interpretations. These claims provide another form of contextual background but unlike evidential claims, they do not require extensive discussion. In general, these doubts emerged when the extension of layer was uncertain for having ambiguous boundaries (Example 3.2) or the extension of a truncated structure was unclear (Example 3.11), but their type and/or existence was not controversial.

Graphic 6.2 illustrates the distribution of observational, evidential and persistent reflexivity in case studies. With these data, it’s very clear then, that despite reflexive diaries included more reflexive information (Graphic 6.1), MRG95 and Tarbat records paid more attention to relevant aspects. In general, evidential reflexivity at MRG95 and Tarbat was triggered by the fragmentariness of remains, hence most reflexive situations had to ponder about the existence or the type of ambiguous features and structures. This doesn’t mean that B49 diggers were constrained to make a relevant reflexive recording, since the preservation of remains was higher at Catalhöyük. However, they had to focus more systematically in the marginal aspects of evidence to ponder about the functional ambiguities of the building. B49 diggers were very successful to revise interpretation of formation process using a feedback mechanism, then it’s very likely that they equally could have used this mechanism to address functional ambiguity more systematically. Evidently, this is not just a task of instruments and mechanisms, because as individual analyses demonstrated, some diggers were more skilled to make relevant reflexive judgments than others. Finally, it’s almost certain that relevant reflexive data in diaries is redundant because this information is already reported in context sheets (see section 3.6). Hence, if a project decides to have a sheets and diary records, they should be designed to cover complementary information. For example, reflexive aspects of features and structures in sheets, and reflexive aspects of the site in diaries.



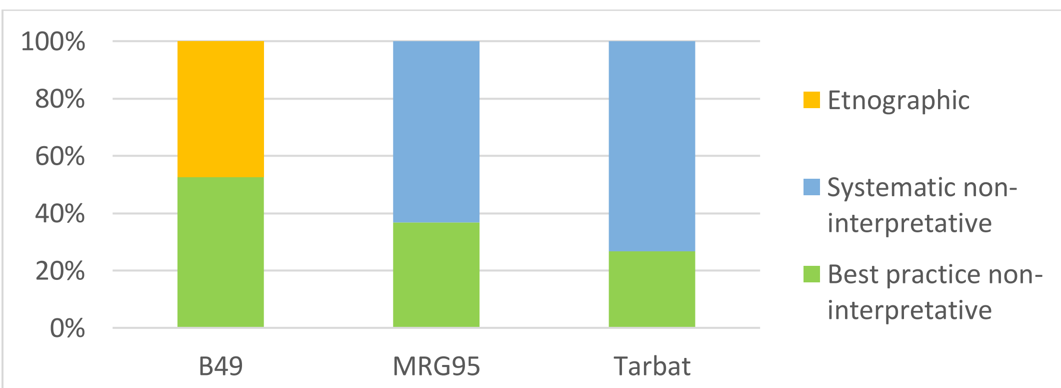
Graphic 6.2. Comparation of reflexive claims by case study. The graphic illustrates the percentage of observational, evidential and persitent claims. Observational claims only include intentional testimonies like narratives of discovery and sketches to represent uncertainty and fluidity. Unintentional testimonies like corrections over stratirgaphic diagrams are not counted.

Finally, Graphic 6.2 indicates the proportion of problematic reflexive testimonies or when evidential testimonies didn't include a discussion. Thereby, the main problem with reflexive diaries is the largest account of irrelevant information whereas the main problems with excavation sheets is the inconsistent quality of relevant reflexive data. Overall, these results demonstrate the need of defining clearer standards for reflexive testimony. First, they must define what sort of interpretive doubts and reconsiderations deserve to be recorded and secondly, they must explain the most adequate format to report such data.

6.3 Decisions and observational circumstances in excavation records

Another common claims of critics of standardized methods is that contexts sheets only include a minimum amount of information regarding technical and observational circumstances. Then, one of the goals of reflexive diaries is providing a better account of these aspects. Particularly, B49 diaries described episodes when observational circumstances affected the process of layer definition and different forms of operative decisions adopted in specific situations (see section 3.5.1). Hence, one of the assumptions of reflexive archaeologists is that these types of testimony would be absent from context sheets.

Context and feature sheets from MRG95 and Tarbat indeed include a systematic component of non-interpretative data that briefly describes the observational condition and the technical procedures implemented onsite (Figure 4.1 and Figure 5.8). Overall, this information is not very interesting except for giving a general idea of these aspects. Nevertheless, there is a smaller number of context sheets that included additional comments to describe distinctive observational challenges and/or operative decisions made when exploring specific layers or features. Among these testimonies, the most interesting describe when observational and operative procedures affected positively or negatively the credibility of interpretations. For example, if the definition of a layer had been compromised by technical errors or if a layer that was barely visible in plan resulted more evident in section. These examples of circumstantial and decision claims in excavation sheets were classified as cases of best practice in non-interpretative data which in general achieve the same function of circumstantial and decision claims in reflexive diaries (Graphic 6.3). At first sight, one could say this evidence suggests that all the systems are equally successful to report this type of non-interpretive data. However, if one considers that the reflexive method includes a sequence of diaries on top of a sequences of excavations sheets, it's reasonable to think that reflexive diaries represent a redundant testimony because context and feature sheets generally report that information already. Finally, it's the issue of ethnographic information in reflexive diaries. A type of non-interpretative information that is highly frequent but irrelevant for contextualizing the credibility of primary data. In sum, standardized sheets are quite successful to report multiple forms of relevant non-interpretative data. Reflexive diaries are equally successful, but evidence indicates that there are problems of redundancy and relevancy with the information they report.



Graphic 6.3. Comparation of non-interpretive claims by case study. Systematic non-interpretative and Best practice non-interpretive categories summarize circumstantial and decision claims by project.

The study also included a brief analysis of sampling testimonies because one of the goals of reflexive diaries is giving account of sampling decisions. However, this is an aspect in which diaries are not very effective. On the contrary, there is a smaller number of context and feature sheets that include explanatory account of sampling decisions (see section 4.5.3 and 5.5.3). In general, most sheets only include a sample number and a brief description of the type of sample. Yet, the rationale of sampling actions always becomes evident after cross-examining this information with interpretative content. Once more, this demonstrates the relevancy of documental analysis to make records speak. Only in a few cases, it was impossible to understand the rationale of sampling activity due the lack of interpretive content in records (see Graphic 4.3 and Graphic 5.3).

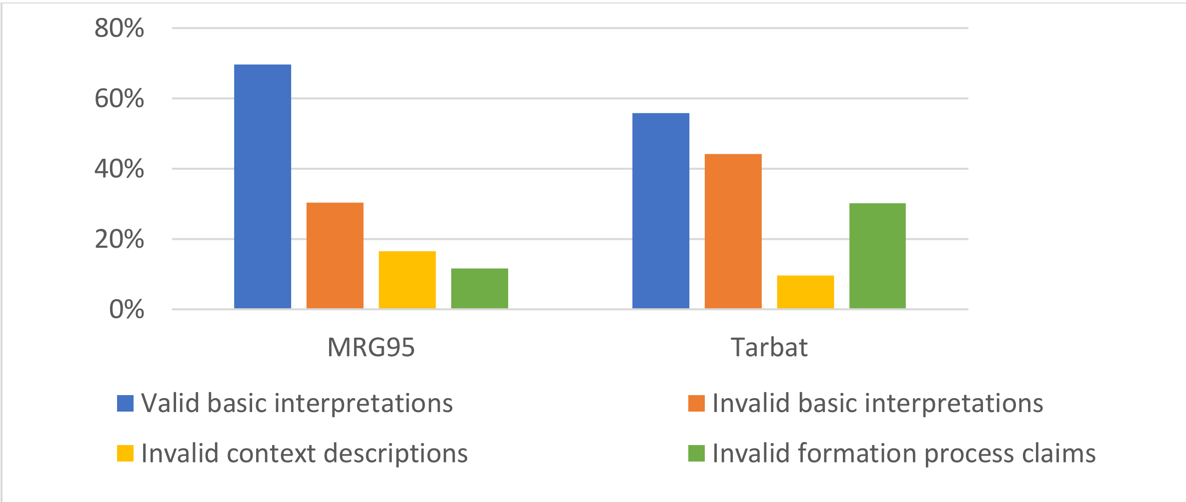
6.4 Credibility of testimonies in excavation sheets

A recurrent belief among reflexive archaeologists is that the interpretive content in standardized formats cannot be evaluated due the lack of an explanatory account of the interpretive process of remains, which apparently is obscured by their predominant descriptive content (Hodder 1999, 31; Mickel 2015, 303; Berggren and Hodder 2003). This claim is groundless because it’s based in a poor understanding of the process of recording and documental analysis. To demonstrate that assessing the validity or credibility of testimonies is possible, a more intensive documental analysis of context and feature sheets from MRG95 and Tarbat was implemented. The analysis was based in different forms of cross-checking, first confronting descriptive and interpretative data and secondly confronting textual and graphic testimonies in context sheets. And occasionally, this also required confronting the content of two or more sheets. This method allowed to identify examples of reliable testimonies and secondly to identify the frequency of reliable and unreliable testimonies.

More importantly, this analysis allowed to identify three recurrent epistemic situations faced by fieldworkers when investigating archaeological remains. First, there are observational situations when the type of layer/feature is evident and records mainly describe the attributes of an observed thing (Example 5.3). Many of these situations frequently arose when reporting well-preserved remains and records usually include a sketch giving testimony of the observed thing (Figure 4.2). Secondly, there are truly interpretative situations when the character of remains is not evident and primary testimonies then must include an inference to explain the interpretative process. Frequently, interpretative testimonies explain depositional processes and occasionally they cover functional aspects. A lot of these testimonies are associated with the investigation of sedimentary layers (Example 4.10 and Figure 4.1) and non-structural features (Example 4.8 and Example 4.13). Nevertheless, interpretive situations can arise too when investigating structures with marginal evidence like floor surfaces (Example 4.3) and postholes (Example 5.4). Overall, interpretative testimonies require more skill to observe and report the interpretation of marginal traces. Finally, there are reflexive testimonies, which describe situations of interpretative doubts and changes that normally require a discussion of evidence to contextualize these claims.

Thereby, observational, interpretative and reflexive situations require different standards of recording to produce reliable testimonies. Hence, when records failed to meet one of these standards, they were considered invalid or unreliable. Graphic 6.4 illustrates the results after a general evaluation of the credibility of testimonies for MRG95 and Tarbat samples. These values have been calculated considering that data from basic claims (layer descriptions, structure and formation processes). As explained before, these claims are closely connected because describing a layer requires defining its type and this indicates its depositional process. The previous internal analysis of MRG95 and Tarbat evaluated the reliability of additional data like features and functional claims. Nevertheless, the credibility of these testimonies derives from the reliability in basic claims (sections 4.5.9 and 5.5.10). Hence, by reducing the evaluation of credibility to basic claims gives a good idea of the quality of samples.

Another important result of this qualitative analysis was identifying the most recurrent types of mistakes. These include invalid layer descriptions, or sheets that describe the attributes of layers, but without stablishing its type (natural, cultural and so forth). Another recurrent problem is when a site record fails to justify formation processes. As explained before, both types of mistake are closely connected because failing to define the type of a layer means failing to define its depositional process. Graphic 6.4 also indicates the number of invalid layer descriptions and invalid formation process claims in each project. Although, it’s pertinent to remind that diggers fail more often to record sedimentary layers and negative features with non-masonry evidence. Whereas, records of structural elements are frequently more reliable. These data coincide with previous observation made elsewhere (Spence 1993, 32). However, no analysis of the type of remains that produced more unreliable sources had been presented before.



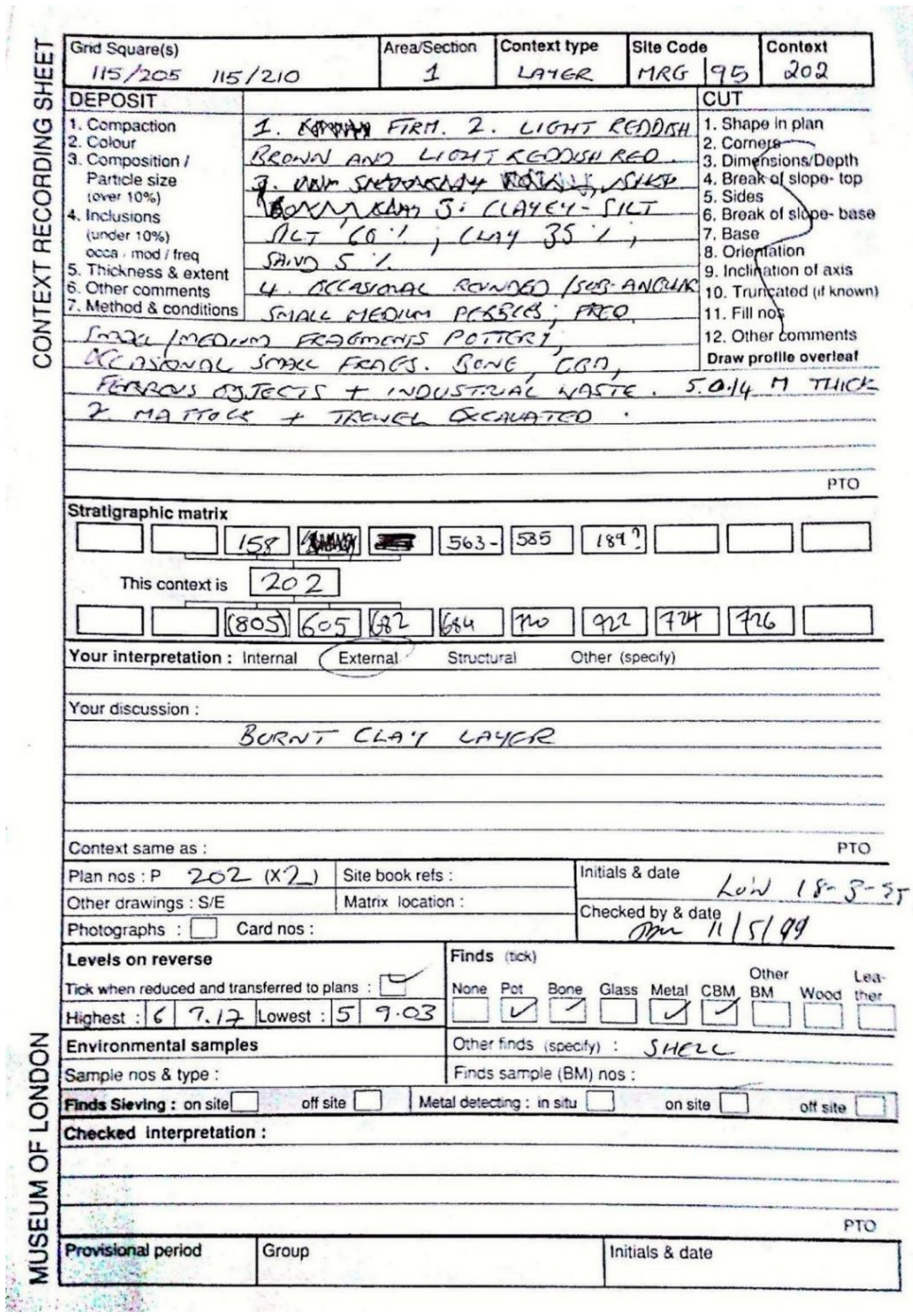
Graphic 6.4. Comparation of the credibility of testimonies by case study. This graphic illustrates the general distribution of valid and invalid testimonies in the total sample of projects. This evaluation is based in the reliability of basic claims (layer description, structural and formation process claims). The graphic also includes the frequency of invalid layer descriptions and formation process claims.

The difference between the credibility of structural and sedimentary layers indicates that observing and interpreting evidence in the last case is more complicated. However, this difference also could be related to specificities in the discovery of sedimentary and structural remains. As explained before, structural remains are frequently exposed (Figure 1.4) whereas sedimentary layers must be observed and recorded and therefore the capture of data might be more complex (Figure 1.5). Hence, this systematic careful recording of structural evidence might be explained by the exposition of vestiges that facilitates the observation and capture of data.

Finally, Graphic 6.4 shows that the percentage of reliable records is higher than the amount of problematic testimonies in both projects, however commercial diggers at MRG95 failed more often to interpret the type of layers, whereas Tarbat diggers failed more often to justify formation processes. Although, a professional workforce was more efficient to produce reliable records generally than a mixed team including professionals and trainees. This evidence coincides with the opinion of a few scholars that subscribe that professional units produce higher quality records than scholar projects (Jones 2013, 181). Moreover, Tarbat records showed very clearly that using a more complex recoding method requires following more rules and the analysis showed clear examples in which trainees have a lot of problems to achieve these tasks, especially for negative features (Figure 5.18).

Finally, there is an additional question regarding the status of unreliable testimonies that at least must be partially tackled. Are these records useless? As explained before, there are some invalid testimonies that could be reinterpreted offsite by the interrogation of descriptive information. Figure 6.1 presents a context sheet from MRG95 that describes a "*burnt clay layer*" (L'ON-202) but failing to stablish an interpretation. However, descriptive information suggest that a firm burned layer of a light reddish-brown colour could be interpreted as a floor. In fact, this context was considered forming part of a larger brickearth surface inside a building (S1) (Seeley and Drummond-Murray 2005, 14) (Figure 4.5). Equally, Figure 6.2 is an example of a context sheet that includes an unjustified interpretation of a fill. However, when confronting the descriptive data of the layer and its associated cut, both suggest the idea of a rubbish pit (Example 5.14). This doesn't imply that every unreliable record can be reinterpreted offsite, as these exercises heavily depends on the quality of descriptions, but in principle problematic records are useful too. Evidently, a more complex re-interpretation of unreliable context and feature sheets is possible by the cross-examination of written and graphic records (photos and drawings). In this way, the value of testimonies not only depends in the reliability of primary data but also in the strategy of documental analysis. In other words, even the most unreliable documents can be sources of knowledge if properly interrogate and even the most reliable sources will be silent in unqualified hands.

Evidently, this should not be an excuse to avoid a preliminary checking of records to highlight errors and inconsistencies onsite (Roskams 2001, 235). The advantage of this procedure has been more evident in Tarbat's records when a more experienced fieldworker occasionally corrects the mistakes of a less experienced digger (Figure 5.19). However, even skilled practitioners can produce problematic testimonies from time to time, because after all, they are human interpreters, not machines. Therefore, having a protocol to check the quality of records among professional diggers should be elementary. Unfortunately, this was not very frequent at MRG95. A recurrent argument in recent years is that such forms of quality control tend to reinforce hierarchical relations in the field, specially between supervisors and diggers (Lucas 2001, 9; Chadwick 2003). In contrast, Roskams argues that supervisors are not always the best option for this task because the sort of revision that they can offer is limited (Roskams 2001, 236). Thereby, a plausible option could be encouraging a sort of peer-review mechanism among diggers or perhaps having a post-excavation officer with more experience in documental analysis. Whatever the case, some form of preliminary checking is necessary.


CONTEXT RECORDING SHEET

Grid Square(s)	Area/Section	Context type	Site Code	Context
115/205 115/210	1	LAYER	MRG 95	202

DEPOSIT
1. Compaction
2. Colour
3. Composition / Particle size (over 10%)
4. Inclusions (under 10%) occa / mod / freq
5. Thickness & extent
6. Other comments
7. Method & conditions

1. FIRM. 2. LIGHT REDDISH BROWN AND LIGHT REDDISH RED. 3. [illegible] SILTY [illegible] 3: CLAYEY-SILT SILT 60 %, CLAY 35 %, SAND 5 %. 4. OCCASIONAL ROUNDED / SUB-ANGULAR SMALL MEDIUM PEBBLES; FREQ. SMALL / MEDIUM FRAGMENTS POTTERY, OCCASIONAL SMALL FRAGS. BONE, CBM, FERROUS OBJECTS + INDUSTRIAL WASTE. 5. 0.14 M THICK 7. MATTOCK + TROWEL EXCAVATED.

CUT
1. Shape in plan
2. Corners
3. Dimensions/Depth
4. Break of slope- top
5. Sides
6. Break of slope- base
7. Base
8. Orientation
9. Inclination of axis
10. Truncated (if known)
11. Fill nos
12. Other comments
Draw profile overleaf

PTO

Stratigraphic matrix
158 | 563- | 585 | 189?
This context is 202
(805) | 605 | 632 | 684 | 710 | 922 | 724 | 726

Your interpretation : Internal External Structural Other (specify)

Your discussion :
BURNT CLAY LAYER

Context same as :
PTO

Plan nos : P 202 (X2) | Site book refs : | Initials & date LOW 18-5-95
Other drawings : S/E | Matrix location : | Checked by & date 11/5/99
Photographs : Card nos :

Levels on reverse
Tick when reduced and transferred to plans : ✓
Highest : 6 7.17 Lowest : 5 9.03

Finds (tick)

None	Pot	Bone	Glass	Metal	CBM	Other BM	Wood	Leather
	✓	✓		✓	✓			

Environmental samples | Other finds (specify) : SHELL
Sample nos & type : | Finds sample (BM) nos :
Finds Sieving : on site off site | Metal detecting : in situ on site off site
Checked interpretation :
PTO

Provisional period	Group	Initials & date

MUSEUM OF LONDON

Figure 6.1. Example of a problematic context sheet that fails to stablish the type of layer being described (LO’N-202).

Y2.0	CONTEXT RECORD	Site: TR99	Context: C1827

Intervention: 24 | Quadrant: Bo

Easting: 871.30 | Northing: 1003.25 | Height: 13.91 m AOD

Recovery level: A B (C) D E

Shape - In plan: IRREGULAR | In profile: DEPOSIT

Description: C1827 WAS IDENTIFIED AT HORIZON 3B AS A VERY DARK BROWN SILTY SAND DEPOSIT, LOCATED IN THE S-E CORNER OF MODULE Bo. INCLUSIONS OF FLECKS OF CHARCOAL, BURNT AND UNBURNT ANIMAL BONES, UNBURNED FISHBONE, BURNT & UNBURNED SANDSTONES AND QUARTZITE PEBBLES, AND SLAG*. THE DEPOSIT MEASURES c. 2.5M IN LENGTH, c. 30 CM IN DEPTH & c. 40-45CM IN WIDTH. THE CONTEXT HAS BEEN CUT BY C1836 TO THE SOUTH, A CONTEXT WHICH AGAIN HAS BEEN CUT BY C1835.

* ALSO SOME INCLUSIONS OF LIMPETS & COCKLES

Condition on recording: DAMP | M.C: DARK REDDISH BROWN | Sieving regime (%): 100

Components from Visual Estimates

%	Distribution	Compaction	Structure	Purity	Colour	Size	Material	A/R/M	D/S/K
49		FRIABLE	GRAINS	CLEAN	5YR 2.5/2	<2mm	SILTY SAND		S
20	DISORDERED	FRIABLE	FLECKS	CLEAN	BLACK	2-5	CHARCOAL		S
15	DISORDERED	HARD	FRAGS	VARIABLE	VARIABLE	20-200	BONES (ANIMAL)		K
1	DISORDERED	HARD	FRAGS	CLEAN	VARIABLE	150-200	BONES (FISH)		K
1	DISORDERED	HARD	FRAGS	CLEAN	VARIABLE	20-50	SLAG		K
1	DISORDERED	HARD	COMPLETE	CLEAN	VARIABLE	20-50	LIMPETS		D
1	DISORDERED	HARD	COMPLETE	CLEAN	VARIABLE	10-20	COCKLES		D
						2 - 6	Gravel		
6	DISORDERED	HARD	COMPLETE	CLEAN	VARIABLE	6 - 20	Gravel	R	D
2	DISORDERED	HARD	FRAGS	VARIABLE	VARIABLE	20 - 60	Pebbles	M	D
3	DISORDERED	HARD	FRAGS	CLEAN	VARIABLE	60 - 200	Pebbles	M	D
1	DISORDERED	HARD	FRAGS	CLEAN	VARIABLE	200 - 600	Cobbles	A	D
						> 600	Cobbles		

Samples: 1 x 10L BULK SAMPLE

Stratigraphy: Over: | Under:
Abuts: | Abutted by:
Same as: | Cut by: C1836

Comments: C1827 HAS BEEN INTERPRETED AS THE BACKFILL OF F345

Belongs to feature: F345 | Identified as: BACKFILL | Status: P / (S)

Date recorded: 25.08.99 | Recorded by: LG

Figure 6.2. Example of an unreliable context sheet that claims the secondary character of the main backfill of F345 but failing to explain the relation between evidence and the formation process (C1827-LG).

6.5 A final contrast. Virtues and problems in recording methods.

The purpose of this investigation was to examine documental evidence to test whether having a more complex recording system improves the quality of primary records and the performance of fieldworkers as it has been suggested (Hodder 2000a, 7; Farid 2000, 25; Carver 2009, 121; Andrews,Barrett and Lewis 2000, 527). However, the comparative analysis demonstrated that the content of primary records is basically the same independently of the recording strategy and the epistemic context of projects. Generally, excavation records contain more descriptive information than interpretative and reflexive data. Still, there are differences in the interpretative content of records, but these are better explained by contextual factors like the character of remains.

Another assumption of reflexive discourse is that standardization strongly determines the content of records (Chadwick 2003, 110; Lucas 2009, 232; Berggren and Hodder 2003, 423). In general, this is true, but documental evidence also demonstrates that descriptions can include attributes not listed in predetermined fields, especially when this data represents important evidence. Likewise, there is some standardization in the way to report an interpretation, but these cannot be uniformized because indicative signs generally vary from case to case. For these reasons, standardized methods are quite flexible. More importantly, the analysis of credibility and the performance of fieldworkers indicated that giving an accurate testimony depends on the skill of fieldworkers to follow standards, but without being constrained by them. Thus, when a fieldworker reports something that escapes the predetermined format, he usually adapts recording rules to his interpretive needs. On the contrary, an untrained fieldworker might not be able to cover even the most basic task that a system demands in the first place. Hence, although a more complex recording strategy might include more spaces for a wider range of data, this doesn't imply that fieldworkers automatically will be able to fulfill these tasks. Overall, this evidence indicates that the performance of fieldworkers and the character of remains are crucial factors in the content and the credibility of records, independently whether data is being recorded in a sheet or an electronic diary.

Evidently, one cannot dismiss the role of methods since the architecture of systems stablishes recording procedures in the first place. However, documental analysis also has shown that methodological flaws can affect the quality of records, for example when a system doesn't stablish very clearly the type of information to be recorded or the guidelines to describe something are ambiguous, thus creating conditions for unreliable, redundant and irrelevant information. Hence, one can set a distinction between the methodological aspects of a recording system or how the architecture of a system contributes or compromises the quality of records for example by stablishing clear or ambiguous protocols for describing a feature. On the other hand, it is the implementation of a recording system in which the performance of fieldworkers and the character of remains contribute or restrain the achievement of interpretative tasks defined by a method. To conclude this contrast, some general observations of the main strengths and failures in the standards and guidelines of recording methods will be summarized. The intention is not to declare a more efficient recording strategy. More exactly, to highlight positive aspects that future projects could imitate and aspects that deserve to be revised or abandoned.

Reflexive diaries

The most important methodological problem of reflexive diaries is the large amount of irrelevant information, which a direct consequence of theoretical problems to justify the purpose of reflexive diaries, especially the narratives of discovery. Another important problem are the various instances of unclear standards in the recording process. For example, some diggers made daily entries, but others made a weekly entry. This also indicated some ambiguities in the goal of records as some entries looked like journal entries and others provided a more synthetic view. Furthermore, relevant content of diaries seems to be highly redundant as much of this data is already captured in standardized formats, especially information about observational circumstances, operative decisions and evidential reflexivity. Similarly, there is repeated information among entries occasionally which also indicates a minor operative problem. Finally, another important problem in reflexive diaries is a heavily dependence on textualization which implies a complete absence of sketches or connections between diaries and graphic materials (drawings, photos) to report relevant matters.

In contrast, one of the strongest aspects of the reflexive method is the regulated interaction among fieldworkers and laboratory specialist in priority tours which contributes for the development of a controlled feedback mechanism with two clear goals. In the first place, priority tours provide a feedback mechanism to exchange previous experience and discuss interpretations among fieldworkers (Example 3.10). Secondly, priority tours create a feedback mechanism between fieldworkers and laboratory staff onsite for revising preliminary interpretations, especially about unobservable aspects like formation processes (Example 3.13). Besides, the testimonies of such exchanges constitute the most relevant data of diaries. This is an important difference with MRG95 and Tarbat where diggers made such kind of retrospective reconsiderations less often (Example 4.11). However, the systematized interactivity at Catalhöyük also had another advantage to check the convergence or inconsistency between independent lines of evidence more immediately. Particularly, this feedback mechanism seems to be an improvement of an early process for 'information retrieval' sketched by David Clarke (Evans,Edmonds and Boreham 2006, 159).

Another positive aspect of reflexive strategies is the explanatory approach especially for spatial and phase data whereas this information was very fragmentary in context sheets. Although, this strategy was not systematically exploded for reasons mentioned above. Thereby, diaries could achieve a similar function to feature sheets that normally include brief explanations of spatial data. In this way, if excavation diaries are going to be implemented alongside other formats, they should function as truly complementary tools.

Single-context recording

One of the most important methodological strengths of the SCR is the use of context sheets to create a sequence of related events to contextualize the interpretation of events and document the fluidity of interpretation. Another important virtue of the system is the standardization of formats for recording a wide range of data for each event. Moreover, the system is sufficiently flexible to give account of observational, interpretative and reflexive situations. Finally, another important virtue of the system is the complementary use to textual and sketches testimonies. The most important problems are some conceptual ambiguities for stablishing types of layers and features. Likewise, the standards for reporting reflexive situations are not very well defined. Finally, there are some problems in the design of the context sheet for the recording of spatial data (Internal/external) (Figure 4.1). Generally, context sheets from MRG95 failed to include an explanatory account of spatial interpretations whereas diaries and feature sheets where more efficient for this task, although these B49 and Tarbat showed a limited use of sketches to report spatial data (see Figure 4.28).

Feature system

The feature system shares most of the virtues of SCR. Furthermore, the feature sheet facilitates the analysis of records by summarizing the interpretation of events into groups. The major methodological problems of the feature system are the lack of a cut sheet and an unclear procedure for the description of features, these problems are derived from ambiguities in the conceptual status of cuts and features. A similar problem with SCR is an ambiguous terminology for defining feature types. Another common problem in both methods is a limited preliminary revision of records onsite which otherwise would have minimized the production of unreliable documents. Otherwise, a systematic preliminary revision of site records would have enabled the development of a formal feedback loop.

6.6 Visual and textual testimony in site records.

Up to this point, issues of recording and interpretation have been discussed without giving much attention to differences between textual and graphic testimonies. This might be an important distinction to make because some archaeologists consider that textual elements have gained primacy lately (Evans 1998a, 189). Partly, this might be consequence of recent depictions of recording as a textual activity (Edgeworth 2003, 2). In fact, some archaeologists consider there is a trend towards textualization (Lucas 2001, 205). Initially, this idea might seem controversial because a modern excavation produces many visual records like plans, sections and photos. However, the notion of textualization can be a plausible idea if understood as an increment in the amount of writing in notebooks and context sheets, and the significance that these devices have gained. For example, early descriptions of SCR focused in the advantages of single-planning and the stratigraphic diagram (Harris 1989, 95). But later, the context sheet became the central element of the system. Moreover, the main data of context sheets, descriptions and interpretation were conceived to be textual (Spence 1993).

In contrast, Chris Evans has explained how antiquarian models functioned as the main elements to give testimony of architecture and proofs of sequences (Evans 2008, 153). Similarly, he has noticed that older archives are primarily visual, like those of Bersu’s explorations in England, which include annotated plans and sections, notebooks with sketches of building reconstructions and photographs and drawings of ethnographic examples of possible similar buildings (Evans 1998a, 185–189). In short, the whole observational and interpretative process is recorded and supported by visual proofs. Similarly, Gavin Lucas has remarked the primary role of illustrations in Pitt-Rivers and Petrie's records (Lucas, 2001, 210). For example, Petrie described the process of recording as planning, copying inscriptions and taking photographs, and only mentions the role of annotations when "*the text is to show the meaning and relation of facts already expressed by form*" (Petrie 1904, 118). Then, although both men have been depicted as opposites due their notions of selective and total record, both considered the visual as the main component of field recording. And not long ago, this standards was similar (Piggott 1965, 166).

Currently, one of the clearest examples of 'textualization' are reflexive diaries which are completely divorced from visual elements. Instead, textual narratives aim to be the back-up of another text, descriptions and interpretations in context sheets. But the information of diaries is easier to understand when one reads the narrative with a graphic reference. Only more recently, reflexive archaeologists have considered the possibility of linking diaries and images with the introduction of 'daily sketches'. These are traced photographs complemented with explanations and interpretations of visual elements, but if we attend the history of diary recording; methodologically there is nothing new, except for the introduction of digital tools (Berggren et al. 2015, 437). Nevertheless, there might be good reasons to conceive some aspects of textualization more positively. In a recent publication, Gavin Lucas presented a typology of archaeological texts which includes descriptions, narratives, explanations and arguments. Although, his main source of study are publications like papers and monographs (Lucas 2019) all of these types of text are found in field records. A context sheet generally includes a description of a layer, and explanation linking evidence to an interpretation, and sometimes it might include a reflexive argument (Figure 4.1). On the other hand, reflexive diaries include narratives and reflexive arguments. This indicates a sophistication of textual recording in recent times that contradicts the general idea that these sources are primarily descriptive.

Interestingly, despite the place of text in recent theoretical discussion in the last decades, recent years have seen an increasing interest in visualization. However, many of these works aim to demonstrate the interpretative nature of drawings and photographs. For example, they argue that illustrations are based in a selection of attributes either when drawing and artefact or mapping the ruins of Teotihuacan (Moser 2014, 97; Olsen et al. 2012, 97; Lucas 2012b, 237). Similarly, these works explain that site photographs are selective for depicting objects from a point of view (Lucas 2001, 211). Hence, a common conclusion of these works is rejecting the notion of an objective visual record since drawings and photos are not copies. Generally, these works acknowledge the evidential value of visual elements, but this is secondary (Olsen et al. 2012, 101; Lucas 2001, 213).

This is an important difference with earlier works. For example, Stuart Piggott and Brian Hope-Taylor acknowledged that illustrations are result of an interpretative process (Hope-Taylor 1966, 109; Piggott 1965, 166), but hey seemed more interested in remarking two main functions of drawings. First, as visual testimonies as "*the drawing enables us to see the soil through the excavator's eye*" (Piggott 1965, 174) and secondly, "*to demonstrate the process of results of interpretation*" (Hope-Taylor 1966, 108). For instance, when a section provides evidence of a sequence. Hence, they understood the selective nature of drawings, but they also knew that chosen attributes were normally selected for their evidential power, thus "*by the very act of realization through drawing, the critical features which will constitute proof of one hypothesis rather than another will be precisely defined*" (Hope Taylor 109; see Moser 2014, 97 for a similar idea). This is an important idea because it implies that selectivity is not arbitrary but guided by a relevancy criterion but more importantly, the visual constituted the main form of recording due its evidential power.

In consequence, the textualization in reflexive diaries implies failing to acknowledge the evidential power of images. Otherwise, a photograph would suffice to prove the existence of B49 and a sequence of images could be the base for capturing the process of discovery. This doesn't mean that text should be vilified as this can be a useful tool to describe obscured aspects in photographs like the constructive materials of structures, which is the core of context sheets indeed. At the same time, one should avoid sanctifying images, as there might be unconvincing illustration like a photograph of the remains of a Mesolithic building (Figure 1.6). Then, text will be useful to highlight a meaningful pattern of postholes, but also to give testimony of data and interpretations not recorded visually. For example, the texture of a layer, likewise there might be some cases when it is more practical transforming visual information into textual data, for example the colour of a layer. In this way, visual and textual proofs can work together for stablishing the credibility of data. MRG95 provided a nice catalogue of multiple ways of using texts and sketches as complementarily testimony (see section 4.6). In fact, modern archaeology has the advantage of being supported by two traditions of archaeological recording. A more recent antiquarian practise that privileged the visual (Evans 1998b, 111), but complemented with an older Greco-Roman tradition of giving historical testimony by written descriptions (ekphrasis) (Ginzburg 2012a).

However, if earlier decades saw a trend towards textualization, recent years have inclined towards visualization and this deserves some attention too. For example, recent techniques like digital planning and photogrammetry have brought an instrumentalization and a specialization of visual recording. Then, diggers are mainly in charge of textual recording and in this context, sketching might be more necessary than ever. However, diggers must be warned to avoid sketching what has been recorded digitally (single plans) and instead encouraged to sketch features and spatial views of the site or perhaps to represents more often interpretative aspects like formation processes (Figure 5.20).

Another important aspect to consider in the future is an over-recording behaviour. For example, since the introduction of digital photography, the number of files in archives has increased exponentially. However, it is not clear whether such extensive recording might be necessary at all. Likewise, the recent trend for 3D site modelling walks in the same direction and a lot of this work is frequently justified in the accuracy and accessibility of data. Interestingly, this view returns to some antiquarian practices of modelling (Evans 1998b), although the difference is obvious. Later versions are created in a virtual environment, but the objective is basically the same, to create a faithful copy (Berggren et al. 2015; Roosevelt et al. 2015; Croix et al. 2019). However, before immersing into such metaphysical issues, one should ask whether such investment of resources is justified. For example, the CRP has produced some 3D models of buildings, but it's not very clear what sort of things can be recorded and proved with a model that cannot be supported with photos and plans. If this model only aims to give testimony of the presence and the form of buildings, it is likely that such investment might be unnecessary, although we must accept that the vividness of 3D models makes them more attractive. However, there are some examples that suggest clearer advantages of 3D modelling. In another example from Catalhöyük, archaeologists explained how 3D modelling from a complex burial area helped to interpret the presence of a cut that was not observed onsite, and this was useful to understand the depositional process of burials (Berggren et al. 2015, 441). Likewise, photogrammetry was crucial to interpret some patterns based in fragmentary evidence at Must Farm in Cambridgeshire (Knight et al. 2019, 651). Hence, it's possible that 3D modelling might be useful to record complex areas for specific problems encountered in the field. Then, it might not be required to reproduce the whole site virtually but instead use this tool more selectively. Evidently, it is possible that archaeologists will find new problems offsite that hadn't been considered before and this would justify modelling the whole site. However, these possibilities must be studied thoroughly.

In sum, archaeologists should consider more seriously the ever-increasing size of archives, either by textualization or visualization. Overall, this is closely connected to the goal of making a total record that stands for the site (Lucas 2012b, 71) and even though this idea has been frequently criticized, even reflexive archaeology ultimately embraces this goal with the production of high definition copies of the site. Additionally, there is a link between the notion of a total record and a having public archive. For example, Chris Evans has mentioned that the idea of a public archive is very recent (personal communication) and most likely it developed from the goal of 'saving the data' for future research. Thereby, the goal of having a total and public archaeological record has become more attainable with digitalization and online storage. Likewise, there is an important moral component as immediate access seems to be an direct route towards democratization and transparency of data (Lukas,Engel and Mazzucato 2018). However, even if these are laudable ideas, it is not very clear whether such massive number of public datasets are the main interest researchers and the general public. Specially, if one considers the investment of time in checking primary records. For instance, beyond internal publications by the CRP and my own research, I know of no other example that has made use of reflexive diaries and the same seems to be true for videos, photographs and excavation sheets. On the contrary, there are many examples of published reports based in different forms of incomplete or problematic backlogged archives (Evans et al. 2016; Evans,Edmonds and Boreham 2006). Hence, it seems plausible that archaeologists should start thinking more seriously about the economy of the archives they produce.

6.7 Individual and collective interpretation in the field.

One of the coincidences between empirical, phenomenological and interpretive approaches is that excavations work is mainly defines as embodied experience or an agent based practice (Edgeworth 2003; Hodder 1997; Lucas 2012a, 242; Roskams 2001, 227; Bender et al. 2007). However, the most relevant discourses acknowledge the incidence of skill, previous experience and length of participation as relevant epistemic factors (Bradley 2003, 155). Based in these premises, an additional analysis was implemented for studying the effects of skill and the length of participation in the individual performance of fieldworkers. Additionally, a longitudinal analysis was implemented to study changes in the interpretative process along the duration of projects.

The results of individual performance demonstrated very clearly the incidence of skill and experience. For example, B49 results showed that less experienced fieldworkers recorded irrelevant information more often whereas professional diggers covered relevant aspects more frequently (Graphic 3.5). On the contrary, results from MRG95 and Tarbat showed that more experienced diggers produced reliable records more often (Graphic 4.6 and Graphic 5.6). Another common characteristic of projects is having frequent changes of team members. For example, Tarbat usually had staff changes between seasons (Table 5.1) but MRG95 had weekly changes (Table 4.1). Evidently, this affects the stability of a team and the level of engagement of participants. Moreover, individual analysis provided evidence that diggers with shorter participation had less chances to develop more complex areas of field investigation, especially spatial, groups of feature and phase interpretations (Graphic 3.5, Graphic 4.7 and Graphic 5.6). Thereby, this evidence indicates that making this type of claims not only depends of the skill of fieldworkers, but also in having an extensive participation in the project that facilitated gaining familiarity with the site and accumulating background knowledge. This interpretation was supported by longitudinal results at B49 and MRG95 that showed how spatial claims became more frequent with the course of time (Graphic 3.8 and Graphic 4.9). Due its excavation strategy, Tarbat showed a more regular production of spatial claims, but the interpretation of building S9 only emerged during the last phase of the project (Graphic 5.8). In short, the more complex interpretative claims require the accumulation of some background knowledge and diggers with longer participation have better chances to do this. Paradoxically, the longitudinal exam also illustrated a decrement in the reliability of testimonies in the final stages of excavations at MRG95 and TRS2, which most likely correlates with time pressures.

Additionally, the analysis showed similar restrictions in recording activities in B49 and Tarbat, where there was a limited participation of trainees in diary recording and feature description. And when they had access, they were not very proficient. This is an important observation considering recent rhetoric to democratize fieldwork. On the contrary, commercial diggers had a more ‘democratic approach’ because there was not differentiated access. Moreover, even though there were differences in the credibility of records, commercial diggers covered the same aspects more systematically, namely descriptive, interpretative and reflexive elements. Thereby, one might suspect that training and systematization are the basis of this behavior. Hence, access to recording activity might be a good start to create a more democratic practice and improve the performance of fieldworkers, but people must be trained, supervised and given better conditions for the development field abilities too.

More generally, the investigation provided multiple examples to observe more clearly the two dimensions of interpretive practice. First, an individual dimension based in agent-based practice and secondly a collective dimension based in the joint action and interactivity among agents. Excavation diaries and sheets are rich sources to illustrate the embodied dimension of interpretive practice, especially when visual and tactile attributes are used to define layers and features, for example (Figure 4.1). Likewise, many testimonies illustrate the relevance of previous experience, for example when making comparative judgments between things being described and things previously observed (see sections 3.5.4, 4.5.4 and 5.5.4).

In relation to collaborative interpretation, reflexive diaries include vivid testimonies of joint action. For example, when describing the discovery of the side room inside B49 (Example 3.14) or the definition of a layer (Example 3.1). Likewise, diaries describe episodes of interactivity among fieldworkers during priority tours (Example 3.10). Excavation sheets are less prodigious to depict joint action and interactivity, but records show the effects of collective work. For instance, a digger initially defines and sketches a layer, and later another digger corrects these testimonies after discovering a correlated layer (Figure 4.17 and Figure 4.18). In another case, a digger defied a timber structure that later was interpreted by another as a cesspit after considering the evidence of a cut (Example 4.14). Overall, these examples describe two forms of coordinated practice. First, there are cases of coordinate practice based in communication or interactivity among agents in which they reach agreements by sharing information and organizing actions. Secondly, there are cases of evidential coordination that emerges when agents react in relation to the actions or information produced by another agent without requiring direct interaction (Weirich 2010, 141). In short, coordination is a necessary condition for collective practice within an epistemic community that must hold common assumptions to be able to share individual experiences across the site (Edgeworth 2003, 47).

Moreover, collective interpretation shows a recurrent pattern in which an individual claim is frequently made in relation to a preliminary framework of facts and beliefs about the site. This produces a collective network derived from the aggregation of individual claims endorsed by a group (List 2011, 224). Such endorsement is possible by giving a level of trust to the performance of agents. However, there is another criterion to endorse new beliefs as these must be consistent with previous beliefs stablished by the group and this constitutes an elementary criteria of collective rationality (Weirich 2010, 139). In short, collective interpretation requires coordination, aggregation and consistency. Then, after considering these attributes, one can even propose that one of the most atypical examples of interpretative practice is a case of collective interpretation. Specifically, the definition of a metalworking area (Figure 5.23) derived from the observations recorded by two fieldworkers that never worked together (ST and NJT) but synthetized in a spatial claim by one of them (NJT) (Example 5.12).

Generally, examples of collective interpretation illustrate cases of belief formation and revision. However, there is one case in which a fieldworker discusses conflicting explanations suggested by two agents (Example 5.13). Likewise, there is plenty evidence showing how agents are self-critic about their own beliefs more often (Example 3.9, Example 4.11 and Example 5.10). The specific reasons of this imbalance are unclear, but it’s possible that agents prefer to adopt a collaborative strategy for the interpretation of the site rather than evaluating competing explanations for two reasons. First, because the task of hypothesis formulation might be more urgent during fieldworks and secondly, because fieldworkers might be less keen to discredit colleagues’ claims. For example, reflexive archaeologists have described some problems for discussing interpretations because staff is reluctant to be questioned and criticised (Berggren and Nilsson 2014, 64; Farid 2000, 25). This coincides with views of socials studies of science that describes the relation between academic status and epistemic trust (Shapin 1994), then by questioning the claims of a team member might be considered a direct attack to his epistemic trust and professional judgment.

In sum, the individual dimension of interpretative practice illustrates the empirical and subjective components of individual performance, but the collective dimension indicates that the work of agents is nourished by the knowledge of a team. In this way, even if diggers act as independent or subjective agents of knowledge, they do not produce subjective views of the site. More exactly, there is a systematic attempt by a network of agents to create a consistent view of the site. The enterprise of individuals creating knowledge is one of the key areas of social epistemology (Goldman 2010, 14). Nevertheless, the principle is already sketched within the empiricist program that describes science as the coordinated labor of many observers (Ladyman 2002, 18).

6.8 Future directions

To conclude this chapter a series of potential themes are considered for future research.

Case studies and methodological appraisal. The selection of case studies was defined by the research agenda of the investigation, but at the time of defining a sampling strategies two options arose. An exhaustive analysis of a few projects or a more superficial analysis of more case studies. In this case, detail was preferred for their relevance for research questions. However, additional case studies would have allowed to compare reflexive practice in commercial and research scenarios for example. Explorations at Terminal 5 for Heathrow airport (Lewis 2006) or in Malmo, Sweden (Berggren 2001, 19) would have been perfect examples for this task. Retrospectively, I think, these examples would have allowed to compare reflexive projects in more challenging scenarios, as the diversity and fragmentariness of remains at Terminal 5 and Malmo seems to be higher and the background knowledge seems more limited than at Catalhöyük.

In the case of MRG95 and Tarbat, these also represent projects in ideal conditions. Hence, it would have been interesting to study ommercial and research investigations in more disadvantageous conditions. For example, a commercial site from a smaller firm or a less researched area than London. Finally, even though there are important differences in the fragmentariness of remains and the deepness of deposits, all the examined sites can be considered urban. Then, it would have been interesting comparing current case studies with a prehistoric investigation. For instance, the most recent explorations at Star Carr could have been a useful case study because this project had an important amount of background knowledge, but the evidence is so fragmentary that most of the interpretations were developed in post-excavation (Milner,Conneller and Taylor 2018).

Additionally, there are further research questions that could be tackled with additional case studies. One the aspects that I'm most interested in considering is the historic dimension of methods. Hence, it would be revealing to compare the results of B49 with diaries with samples from earlier and later phases of the project. This would allow to observe continuities and changes in reflexive methods over a period of 25 years. For example, this contrast should provide evidence of the effects of the introduction of digital tools. Similarly, a contrast between commercial archives from the Rescue, the PPG16 and the NPPF eras would be crucial to observe modifications in the format of context sheets and examine whether these changes might affect recording process over a long period. Finally, it would be important to have a historical but also a geographic analysis of excavation diaries to compare formats of recording and the type of content. I think this would provide a richer picture of the way that archaeologists have used these tools.

<u>Documental analysis and archives</u>. This is another area that had some methodological limitations due its focus on textual formats. Hence, future research could give more emphasis to the relation between textual and graphic testimonies, specially to tackle issues of credibility. Hence, instead of paying much attention to the relation between context sheet and excavation diaries, documental analysis could be reformulated as checking context sheets or diaries against photos and drawings. This procedure should bring more light about the process of identifying reliable and unreliable records, but also to develop an analytical strategy to reinterpret unreliable sources.

Additionally, other research questions could be tackled by studying different types of archives and paying more attention to the relation between textual and graphic testimonies. People from the Cambridge Archaeological Unit have provided many examples of distinct types of analysis of backlogged archives making clear examples of the importance of confronting textual and graphic sources. However, these ideas could be enriched by considering differences between examining an archive with a clearer research agenda (Evans,Edmonds and Boreham 2006) and another with more ambiguous objectives (Evans et al. 2016) or to compare the exercise of reinterpreting an archive based in known procedures with an archive where standards are unfamiliar (Evans and Appleby 2008). This type of studies should provide more tangible evidence against another common theoretical premise, that records created under no research agenda or with a different agenda to ours are not useful (Binford 1964; Reynolds and Barber 1984; Carver 2011).

<u>Experiments with recording methods</u>. One of the most immediate areas of opportunity defined by this investigation is the possibility to implement controlled experiments for improving formal aspects and guidelines. For example, modify the format of context sheets, especially in sections for spatial data or introduce corrections in the design of features sheets for improving feature descriptions. Evidently, this isn't the first time that formal revision is proposed (Chadwick 2003, 108; Adams 2000). However, former attempts suggested radical changes, whereas my view is more conservative and only aims for small changes to fix punctual problems. Another possibility is improving the instructions for filling up context sheets, particularly to emphasise differences when recording well preserved and fragmentary remains. Likewise, remark the importance of giving equal attention to sedimentary and structural layers. Similar experiments could be run with excavation diaries to emphasize the need of focusing in higher-order aspects. Evidently, these experimental try-outs will not improve automatically as the quality of records depends in the skills of fieldworkers, then as many people has recommended already, training schemes and job-opportunities should be improved too (Everill 2009; Edgeworth 2011). Altogether, this would represent an important step towards the development of an experienced workforce and paving the way towards a social context in which archaeological knowledge is not only conceived in terms of theoretical discourse, but in terms of practical work.

<u>Visualization and instrumentalization</u>. Another potential subject of study is detailing the criterial for textual and visual recording. This investigation should clarify the potential use of these forms of representation, specially to refine the awareness in which a graphic testimony should have the primacy and vice versa. Another subject of further study should be the relation between visualization and instrumentalization, particularly to investigate the consequences of the specialization of digital visual records and the enlargement of the archive as mentioned before (see section 6.6).

<u>Theoretical and experimental work on individual and collective interpretation</u>. Finally, another possibility is expanding theoretical notions of interpretative fieldwork as individual and collective enterprise. This revision could help to define experimental approaches to field practice. Evidently, this field has been already explored by reflexive archaeology. However, a more detailed reflection could bring more interesting changes to recording practice. For example, context and feature description could be mainly defined as individual forms of field recording (even though this is not always the case). Meanwhile, diary recording could be reformulated as a

collective enterprise in which the views of a team could be summarized to advance in the interpretation of the site, instead of each agent recording independent diary entries. Another potential area of study could be investigating in more detail the circumstances in which fieldworkers follow a collaborative behaviour to explain something and those in which they adopt a competitive strategy, for example when fieldworkers do not agree in the interpretation of something.

Concluding remarks

To conclude, I will return to the starting point, namely, the relation between interpretive theory and methodological discourse but from a different angle. In the course of this investigation, it has been shown that methodological discourses pay a lot of attention to standardization and instrumentalization. Standardization is elementary in the principles of SCR as depicted by the Museum of London because it encourages systematization of practices, whereas instrumentalization is crucial for reflexivity because it promotes the use of technology, namely excavation diaries, feature sheets, video cameras, tablets and soon. Evidently, both ideals should be welcomed in principle. Nevertheless, SCR and reflexivity adopt different attitudes regarding the role of fieldworkers in the creation of knowledge. Specifically, the Museum of London gives equal weight to standardization and the performance of fieldworkers for the quality of data. Reflexivity is unclear about standardization (even though the best examples of reflexive practice are directly dependent on it), but they firmly believe that fieldworkers using complementary tools will perform much better, independently of their expertise and training. For this reason, reflexivity puts a lot of confidence in technology. The problem with this idea is that evidence indicates that only the best trained diggers can make the best use of instruments. Thus, as final reflection, I want to consider other forms of instrumentalism that permeate methodological discourse and how this contributes to minimize the role of agents and the interpretative nature of archaeology.

Previously, I have mentioned how British archaeologists have aimed to improve the performance of fieldworkers by radically changing the design of context sheets (Chadwick 2003, 108; Adams 2000). These attempts include moderate to strong versions of instrumentalism depending on whether they consider the performance of fieldworkers relevant. Moreover, these examples are interesting because despite there is an important theoretical understanding about field interpretation, British archaeologists have a strong commitment with the idea that if you provide fieldworkers with a slightly different tool, they will perform completely different. Interestingly, a stronger version of this idea has emerged beyond the islands, notably in countries that have imported the context sheet in the last thirty years. British stratigraphic methods started to become popular in Europe and America in the nineties (Harris,Brown and Brown 1993) and more recently, SCR has expanded to Nordic countries, some areas of Germany, France, Italy and Spain where local versions of the context sheet have been adapted (Pavel 2010a). Furthermore, there have been attempts to introduce SCR in Egypt (Tassie 2015) and Eastern Europe (Pavel 2010b, 28). Generally, the main justification for substituting local methods (commonly the arbitrary and the 'box-system') is the superiority of stratigraphic procedures as argued by Harris long ago. Nevertheless, most of these attempts are not very clear about aspects like the training and the performance of fieldworkers. And this should be a crucial matter if one considers that in many of those places the tradition of separating digging and recording activities still continues. For example, Richard Greatorex has described some failed attempts to introduce SCR in Germany, because diggers were reluctant or insufficiently trained for recording work, and then supervisors had to cover this task (Greatorex 2004, 266). This example replicates some aspects of my early experience as an archaeology student in Mexico, where I was trained to use SCR. However, when I got my first contract in Mesoamerican archaeology within a project that apparently followed this method, most of my job consisted in supervising workers defining layers that later I had to describe in context sheets. Again, digging and recording activities were separated as it's generally the case in Mesoamerican archaeology. Then, there was some unjustified pride, because we thought we were producing higher quality records than the average Mesoamerican project. Certainty, this is very limited evidence, but there was a faith in standardized formats, without considering the important matters of who and how these sheets are filled. Hence, this indirectly minimizes relevant aspects about the performance of agents. Hence, one of the greatest virtues of MOLA's rhetoric is the importance they assign to these aspects.

Reflexivity is a more recent form of instrumentalization that has praised the role of technology since its earliest years, when they recommended the use of electronic diaries and video recording (Hodder 1997, 699, 1999, 117). More recently, the emphasis has changed towards the use of digital tools like tablets and 3D recording methods like photogrammetry and laser-scanning (Berggren et al. 2015; Taylor et al. 2018). In the case of digital tablets, there is greater awareness of the role of fieldworkers because they must be trained to use this technology. However, the introduction of f tablets has raised concerns whether digital tools might produce disengagement or loss of interpretative skills (Morgan and Wright 2018, 2). However, some supporters of tablet technology have correctly claimed that many archaeologists falsely assume that paper records are intrinsically better without considering that these were frequently recorded by trainees (Ellis 2016, 60). Hence, promoters of tablets argue that recording methods are basically the same but with a digital tool (Roosevelt et al. 2015, 339). So far, one of the clearest examples of disengagement comes from Catalhöyük as digital planning requires taking a photo that later is vectorized in a tablet, then many fieldworkers have been caught 'recording' outside the trenches (Taylor et al. 2018). However, this shouldn't give the false impression that paper records automatically create an inclination of making an engaged record, as we all know, cases of filling context sheets in the sides of the trench or far worst when the layer has been removed are not uncommon. A more controversial example is when tablet endorsers claim that these devices can improve the interpretative process of fieldworkers by having immediate access to previous data and reports (Berggren et al. 2015, 443). Although, this is a claim remains to be tested, it is very telling because it is a clear attempt to substitute the progressive accumulation of background knowledge with an instrument. In general, the most obvious advantages of tablet recording are improvements in the capture, accessibility and storage of data. In contrast, improvements on the quality of data seem very limited and the clearest example is a reduction of illegible information in textual records. Hence, many of these discourses promote the false notion that simplifying recording work is the same as improving the quality of data. For this reason, it's exaggerated when some promoters of tablet technology claim: "*the iPad thus radically transformed the ways in which we recorded, and engaged with, the excavation of a large urban site*" (Ellis 2016, 56). Then, even if these discourses do not ignore the role of agents, they certainly place great confidence in tools.

In relation to 3D-recording methods, supporters mention improvements in accuracy, accessibility and the efficiency to capture visual data (Roosevelt et al. 2015, 326; Berggren et al. 2015, 446; Croix et al. 2019). In this case, the high-resolution of visual data is the most obvious improvement in the quality of records. For example, archaeologists working at Catalhöyük have reported some experiments of laser scanning of Neolithic buildings (Berggren, 2015, 437) and a team of American archaeologists have provided examples of photogrammetry in Bronze age structures (Roosevelt et al. 2015, 337). However, as mentioned before, one of the most immediate consequences of this technology is that diggers are excluded from visual recording because these methods require a level of specialization (Sapirstein and Murray 2017, 337). Moreover, these examples seem to celebrate obscuring the presence of fieldworkers, by removing the marks of the interpretative process in visual records, namely the selective nature of sketches and photos.

One of the clearest examples in which 3D-modelling is used as an attempt to minimize the presence of fieldworkers comes from a recent Nordic project that systematized the laser-scanning of each layer. In their view, this is a better method because it creates a high-resolution model of layers, and therefore this reduces the subjectivities and imperfections of hand-made drawings. Moreover, they point out that having a 3D model captures all the physical contacts and the undulations of surfaces that enables further assessment of the sequence and formation processes in post-excavation. Then, this form of instrumentalism not only places greater confidence in tools, furthermore, it conceives fieldworkers as unreliable sources of knowledge. One of the most astonishing aspects of these views is their resemblance with nineteen century objectivity programs that aim to restrain the agency of observers because they could risk the process of gaining knowledge (Daston and Galison 2007). In fact, the Nordic project is the very clear about it, when they claim: "*One might assume that technological and methodological innovations such as 3D recording, and micromorphology bring ideal objectivity to ever closer*" (Croix et al. 2019). Evidently, this is very hard to conceive because layers, features and structures recorded in high-resolution are defined by fieldworkers in the first place. For this reason, once more, there is an overreaction when promoters of 3D methods, qualify these technological changes in the 'digital turn' (Taylor et al. 2018) as a 'paradigm shift' (Roosevelt et al. 2015, 339); particularly, when the goals and methods of this technology are renewed versions of objective science. Similarly, having high-resolution models doesn't imply that interpretation will be brought to trowel's edge more immediately because a 3D model of buildings doesn't imply having a better understanding of the functional aspects of a site. Thereby, archaeologists must think more carefully whether instrumental changes, indeed represent methodological innovations, because the second doesn't necessarily follow from the first. Perhaps, this is a disciplinary bias as quite often archaeologists interpret evidence of technological change as cultural revolution. However, if we attend the history of the disciple, for example the stratigraphic revolution of the seventies in Britain was not triggered by an important technological change, instead this methodological revolution was grounded in a new way to dissect and record sites which ultimately led to technological change with the invention of stratigraphic diagrams and context sheets. In other words, instrumental change was a consequence not a cause. In this way, we can also claim that the most valuable elements of reflexive method don't come from instrumental innovations but from the refinement of feedback mechanism.

In this landscape, MOLA's school deserves more credit than what traditionally has received because by putting equal trust in methods and the trained judgment of fieldworkers, they embrace the principles of an interpretive science. Moreover, the level of specialization that British fieldworkers have achieved is one of the necessary conditions that enables the implementation of reflexive methods. If this sounds too far-fetched, one should consider whether reflexive methods still makes sense when a labourer is being told what to do by a not very experienced archaeologists taking notes from the side of the trench (Leighton 2015, 84). Equally interesting is noticing the influence of British methods overseas. In this process, it's possible that British archaeologists have failed to communicate that SCR and reflexivity are not simply about using context sheets, reflexive diaries and so on. Then, recipients of British methodology fail to capture that the biggest asset of British Archaeology is its highly specialized workforce (Roskams 2001, 27). However, this is something that British archaeologists fail to acknowledge quite often too. Hence, despite theoretical work has provided a clearer picture of the interpretive nature of fieldwork, the British community maintains a methodological view in which the quality of data is mainly conceived as the result of improvements in machinery and methods. Then, if the discipline really wants to embrace its interpretative nature, we need to strength the bound between interpretative theory and methodology to give equal value to the performance of fieldworkers, tools and procedures. Evidently, technological changes must be welcomed. For example, improvements in connectivity and networking should facilitate the operation and update of mechanism for collective interpretation and feedbacking, for instance live streamed tours to interact with site/period specialists more frequently. But we don't require fancy theory to justify the introduction of a more efficient tool. Similarly, online recording should facilitate checking-up records, but without assuming, that the quality of data is the consequence of machinery. Finally, if the discipline truly accepts its interpretative nature, this might be reflected in the laboural conditions of fieldworkers. Hence, they will be considered an investment rather than an expense. And perhaps, projects might not be only interested in having the best technology onsite, moreover they will be interested in having the best trained archaeologists in the field. Hopefully, this investigation might have done something useful to achieve these tasks.

References

Adams, M. (2000). The optician's trick: an approach to recording excavation using an iconic formation process recognition system. In: S. Roskams (Ed). *Interpreting Stratigraphy: Site evaluation, recording procedures and stratigraphic analysis.* Archaeopress. pp.91–101.

Adams, M. and Brooke, C. (1995). Unmanaging the past: Truth, data and the human being. *Norwegian Archaeological Review*, 28 (2), pp.93–104. [Online]. Available at: doi:10.1080/00293652.1995.9965587.

Aitchison, K. (2009). After the 'gold rush': global archaeology in 2009. *World Archaeology*, 41 (4), pp.659–671. [Online]. Available at: doi:10.1080/00438240903363772.

Aliseda, A. (2004). Logics in scientific discovery. *Foundations of Science*, 9 (3), pp.339–363.

Aliseda, A. (2005). The Logic of Abduction in the Light of Peirce's Pragmatism. *Semiotica*, 4 (153), pp.363–374.

Allchin, D. (2001). Error types. *Perspectives on Science*, 9 (1), pp.38–58. [Online]. Available at: doi:10.1162/10636140152947786.

Andrews, G., Barrett, J. and Lewis, J. (2000). Interpretation not record: The practice of archaeology. *Antiquity*, 74 (285), pp.525–530. [Online]. Available at: doi:10.1017/S0003598X00059871.

Austin, J. L. (1962). *Sense and sensibilia*. Warnock, G. (Ed). Oxford : Clarendon Press.

Baker, E. (2011). The Manpower Services Commission and La Grava. In: Schofield, J. (Ed). *Great Excavations*. Shaping the Archaeological Profession. Oxbow Books. pp.194–211. [Online]. Available at: http://www.jstor.org/stable/j.ctt1cd0nz2.18.

Barker, P. (1969). Some aspects of the excavation of timber buildings. *World Archaeology*, 1 (2), pp.220–235. [Online]. Available at: doi:10.1080/00438243.1969.9979441.

Barker, P. (1982). *Techniques of archaeological excavation*. 2nd ed. New York : Universe Books.

Barrett, J. (2013). Archaeologically Sustainable Development in an Urban Context. *Papers from the Institute of Archaeology*, 23 (1), pp.1–4. [Online]. Available at: doi:http://doi.org/10.5334/pia.438.

Barrett, J. (2016). Archaeology after interpretation. Returning humanity to archaeological theory. *Archaeological Dialogues*, 23 (2), pp.133–137. [Online]. Available at: doi:10.1017/S1380203816000167.

Barrett, J., Bradley, R. and Hall, M. (1991). *Papers on the prehistoric archaeology of Cranborne Chase*. Oxford : Oxbow Books.

Bender, B. et al. (2007). *Stone worlds : narrative and reflexivity in landscape archaeology*. Hamilton 1954-, S. et al. (Eds). Walnut Creek, CA : Left Coast Press.

Berggren, A. (2015). Reflexive approaches in archaeology. In: Claire Smith (Ed). *Encyclopedia* *of Global Archaeology*. Springer. pp.6249–6258. [Online]. Available at: doi:10.1007/978-3-319-51726-1.

Berggren, Å. (2001). Swedish Archaeology in Perspective and the Possibility of Reflexivity. *Current Swedish Archaeology*, 9, pp.9–24.

Berggren, Å. et al. (2015). Revisiting reflexive archaeology at Çatalhöyük: integrating digital and 3D technologies at the trowel's edge. *Antiquity*, 89 (344), pp.433–448. [Online]. Available at: doi:10.15184/aqy.2014.43.

Berggren, A. and Hodder, I. (2003). Social Practice , Method , and Some Problems of Field Archaeology. *American Antiquity*, 68 (3), pp.421–434.

Berggren, A. and Nilsson, B. (2014). Going back, looking forward: reflexive archaeology or reflexive method? In: I. Hodder (Ed). *Integrating Çatalhöyük : themes from the 2000-2008 seasons*. British Institute at Ankara. pp.57–70.

Binford, L. R. (1964). A Consideration of Archaeological Research Design. *American Antiquity*, 29 (4), pp.425–441. [Online]. Available at: doi:10.2307/277978.

Bloch, M. (1954). *The historian's craft*. Manchester University Press.

Bowsher, J. and Miller, P. (2009). *The Rose and the Globe playhouses of Shakespeare's Bankside, Southwark: excavations 1988-91*. London : Museum of London Archaeology.

Bradley, R. (1997). To see is to have seen. In: B. L. Molyneaux (Ed). *The cultural life of images: visual representation in archaeology*. Routledge. pp.62–72.

Bradley, R. (2003). Seeing things. Perception, experienceand the constrains of excavation. *Journal of Social Archaeology*, 3 (2), pp.151–168.

Bradley, R. (2006). Bridging the Two Cultures. Commercial Archaeology and the Study of Prehistoric Britain. 2008/11/28. *The Antiquaries Journal*, 86, pp.1–13. [Online]. Available at: doi:DOI: 10.1017/S0003581500000032.

Bradley, R. and Edmonds, M. (1993). *Interpreting the axe trade : production and exchange in Neolithic Britain*. Cambridge .

Brand, G. (1979). *The central texts of Ludwig Wittgenstein*. Oxford : Blackwell.

Carver, G. (2012). How to archaeologize with a hammer. In: Cobb, H. et al. (Eds). *Reconsidering Archaeological Fieldwork: Exploring On-Site Relationships between Theory and Practice*. Springer. pp.15–29. [Online]. Available at: doi:10.1007/978-1-4614-2338-6.

Carver, M. (1989). Digging for ideas. *Antiquity*, 63 (241), pp.666–674. [Online]. Available at: doi:10.1017/S0003598X00076808.

Carver, M. (2005). *Sutton Hoo: a seventh-century princely burial ground and its context*. London : British Museum Press.

Carver, M. (2009). *Archaeological investigation*. London : Routledge.

Carver, M. (2011). *Making archaeology happen: design versus dogma*. California : Left Coast Press.

Carver, M. (2015). Field method in archaeology: an overview. In: M. Carver, B. Gaydarska and S. Monton-Subias (Eds). *Field archaeology from around the world*. Springer. pp.3–9.

Carver, M. (2016). *Portmahomack: monastery of the Picts*. 2nd ed. Edinburgh University Press.

Carver, M., Garner-Lahire, J. and Spall, C. (2016). *Portmahomack on Tarbat Ness: changing ideologies in North-east Scotland, sixth to sixteenth century AD*. Society of Antiquaries of Scotland.

Carver, M. O. H. (1987). *Underneath English towns : interpreting urban archaeology*. London : B.T. Batsford.

Chadwick, A. (1998). Archaeology at the edge of chaos: further towards reflexive excavation methodologies. *Assemblage*, 3 (3), pp.1–6. [Online]. Available at: https://archaeologydataservice.ac.uk/archives/view/assemblage/html/3/3chad.html.

Chadwick, A. (2003). Post-processualism, professionalization and archaeological methodologies. Towards reflective and radical practice. *Archaeological Dialogues*, 10 (1), pp.97–117. [Online]. Available at: doi:10.1017/S1380203803001107.

Childe, G. (1956). *Piecing together the past*. London : Routledge.

Clark, G. (1954). *Excavations at Star Carr: an early mesolithic site at Seamer, near Scarborough, Yorkshire*. Cambridge University Press.

Clark, G. (1957). *Archaeology and society: reconstructing the prehistoric past*. London : Methuen.

Clark, P. (1992). Contrast in the recording and interpretation of rural and urban stratifications. In: *Interpretation of stratigraphy: a review of the art . Proceedings of a conference held on 18th June, 1992, at the City of Lincoln*. City of Lincoln Archaeology Unit. pp.17–19.

Conneller, C. et al. (2012). Substantial settlement in the european early mesolithic: New research at Star Carr. *Antiquity*, 86 (334), pp.1004–1020. [Online]. Available at: doi:10.1017/S0003598X00048213.

Croix, S. et al. (2019). Single Context, Metacontext, and High Definition Archaeology: Integrating New Standards of Stratigraphic Excavation and Recording. *Journal of Archaeological Method and Theory*, pp.1–41. [Online]. Available at: doi:10.1007/s10816-019-09417-x.

CRP. (2014). *ÇATALHÖYÜK : CRIB SHEET*. Catalhoyuk Research Project.

Daston, L. and Galison, P. (2007). *Objectivity*. Zone Books.

Demoule, J.-P. (2002). Rescue archaeology. The French way. *Public Archaeology*, 2 (3), pp.170–177.

Demoule, J.-P. (2011). We still have to excavate – but not at any price. 2011/04/21. *Archaeological Dialogues*, 18 (1), Cambridge University Press., pp.5–10. [Online]. Available at: doi:DOI: 10.1017/S1380203811000043.

Department of Physics University of Cambridge. (2014). *Keeping Laboratory Notes and Writing*

Formal Reports.

Eddisford, D. (2006). Building 49. In: *Catalhoyuk 2006 archive report*. pp.71–78. [Online]. Available at: http://www.catalhoyuk.com/archive_reports/2006.

Eddisford, D. (2008). Building 49 Archive report 2008. In: *Catalhöyük 2008 archive report*. Catalhoyuk Research Project. pp.30–38. [Online]. Available at: http://www.catalhoyuk.com/archive_reports/2008.

Eddisford, D. (2014). Chapter 14. Building 49. In: I. Hodder (Ed). *Çatalhöyük excavations: the 2000-2008 seasons: the 2000-2008 seasons*. pp.313–356.

Eddisford, D. and Morgan, C. (2019). Single Context Archaeology as Anarchist Praxis. *Journal of Contemporary Archaeology*, 5 (2), pp.245–254. [Online]. Available at: doi:10.1558/jca.33580.

Edgeworth, M. (2003). *Acts of discovery: an ethnography of archaeological practice*. Oxford : Archaeopress.

Edgeworth, M. (2011). Excavation as a ground of archaeological knowledge. *Archaeological Dialogues*, 18 (1), Cambridge University Press., pp.44–46. [Online]. Available at: doi:DOI: 10.1017/S1380203811000109.

Edgeworth, M. (2012). Follow the Cut, Follow the Rhythm, Follow the Material. *Norwegian Archaeological Review*, 45 (1), pp.76–92. [Online]. Available at: doi:10.1080/00293652.2012.669995.

Ellis, S. J. R. (2016). Are We Ready for New (Digital) Ways to Record Archaeological Fieldwork ? A Case Study from Pompeii. In: A. Erin-Walcek, Gordon, J. M. and D. Counts (Eds). *Mobilizing the Past for a Digital Future*. Grand Forks : The Digital Press at the University of North Dakota. pp.51–75. [Online]. Available at: https://thedigitalpress.org/mobilizing-the-past-for-a-digital-future/.

Embree, L. (1992). Phenomenology of a Change in Archaeological Observation. In: L. Embree (Ed). *Metaarchaeology*. Kluwer academic publisher. pp.165–193. [Online]. Available at: doi:10.1007/978-94-011-1826-2_7.

English Heritage. (1991). *Management of archaeological projects*. London : English Heritage.

English Heritage. (2002). *A research framework for London archaeology 2002*. T. Nixon (Ed). London : Museum of London.

Enloe, J. G. and Audouze, F. (2010). The Magdalenian Site of Verberie (Le Buisson Campin). An overview. In: E. Zubrow, J. Enloe and F. Audouze (Eds). *The Magdalenian household: unraveling domesticity*. State University of New York Press. pp.15–21.

Evans, C. (1989). Archaeology and modern times: Bersu’s Woodbury 1938 & 1939. 2015/01/02. *Antiquity*, 63 (240), Cambridge University Press., pp.436–450. [Online]. Available at: doi:DOI: 10.1017/S0003598X00076419.

Evans, C. (1998a). Constructing Houses and Building Context: Bersu’s Manx Round-house Campaign. 2014/02/18. *Proceedings of the Prehistoric Society*, 64, Cambridge University Press., pp.183–201. [Online]. Available at: doi:DOI: 10.1017/S0079497X00002218.

Evans, C. (1998b). Modelling monuments and excavations. In: Chadarevian, S. de and Hopwood, N. (Eds). *Models. The third dimension of science*. Stanford University Press. pp.109–137.

Evans, C. (2008). Model excavations: ‘Performance’ and the three-dimensional display of knowledge. *Archives, Ancestors, Practices: Archaeology in the Light of its History*, pp.147–161.

Evans, C. et al. (2016). *Lives in land: Mucking excavations by Margaret and Tom Jones, 1965-1978. Prehistory, context and summary*. Oxbow Books.

Evans, C. and Appleby, G. (2008). Historiography and Fieldwork: Wyman Abbott’s Great Fengate Ring-Ditch (a Lost Manuscript Found). *Proceedings of the Prehistoric Society*, 74, pp.171–192. [Online]. Available at: doi:DOI: 10.1017/S0079497X00079342.

Evans, C., Edmonds, M. and Boreham, S. (2006). ‘Total Archaeology’ and model landscapes: Excavation of the Great Wilbraham causewayed enclosure, Cambridgeshire, 1975-6. *Proceedings of the Prehistoric Society*, 72.

Everill, P. (2009). *The invisible diggers: a study of British commercial archaeology*. Heritage Marketing & Publications Ltd.

Everill, P. (2010). British Commercial Archaeology: Antiquarians and labourers; developers and diggers. In: Y. Hamilakis and P. Duke (Eds). *Archaeology and Capitalism*. Walnut Creek, CA, USA : Left Coast Press. pp.119–136.

Farid, S. (2000). The excavation process at Catalhoyuk. In: I. Hodder (Ed). *Towards reflexive method in archaeology: the example at Çatalhöyük*. McDonald Institute for Archaeological Research, University of Cambridge. pp.19–36.

Farid, S. (2015). Proportional representation: multiple voices in archaeological interpretation at Catalhoyuk. In: R.Chapman and A. Wylie (Eds). *Material Evidence Learning from Archaeological Practice*. Routledge. pp.59–78.

Flannery, K. V. (1976). *The Early Mesoamerican village*. New York : Academic Press.

Flannery, K. V. (1982). The Golden Marshalltown: A Parable for the Archeology of the 1980s. *American Anthropologist*, 84 (2), pp.265–278. [Online]. Available at: doi:10.1525/aa.1982.84.2.02a00010.

Garrow, D. (2012). Odd deposits and average practice. A critical history of the concept of structured deposition. 2012/11/26. *Archaeological Dialogues*, 19 (2), pp.85–115. [Online]. Available at: doi:DOI: 10.1017/S1380203812000141.

Ginzburg, C. (1989). Clues: roots of an evidential paradigm. In: C. Ginzburg (Ed). *Clues, myths and the historical method*. Johns Hopkins University Press. pp.87–113.

Ginzburg, C. (1991). Checking the Evidence: The Judge and the Historian. *Critical Inquiry*, 18 (1), pp.79–92. [Online]. Available at: doi:10.1086/448624.

Ginzburg, C. (1999). *The judge and the historian : marginal notes on a late-twentieth-century miscarriage of justice*. London : Verso Books.

Ginzburg, C. (2012a). Description and Citation. In: *Threads and Traces*. True False Fictive. 1st ed. University of California Press. pp.7–24. [Online]. Available at: http://www.jstor.org/stable/10.1525/j.ctt1ppqkr.5.

Ginzburg, C. (2012b). Microhistory: two or three things that I know about it. In: *Threads and Traces: True, flalse, fictive*. University of California Press. pp.193–214.

Ginzburg, C. (2012c). *Threads and traces : true, false, fictive*. Berkeley : University of California Press.

Ginzburg, C. (2018). *Cinco reflexiones sobre Marc Bloch*. Mexico : Contrahistorias.

Goldman, A. I. (2010). A guide to social epistemology. In: A. Goldman and D. Whitcomb (Eds). *Social Epistemology: Essential Readings*. Oxford University Press. pp.11–37.

Grayson, D. K. and Meltzer, D. J. (2002). Clovis hunting and large mammal extinction: A critical review of the evidence. *Journal of World Prehistory*, 16 (4), pp.313–360. [Online]. Available at: doi:10.1023/A:1022912030020.

Greatorex, R. (2004). Towards an integrated european recording system for archaeology: The problems that lie ahead. In: G. Carver (Ed). *Digging in the dirt. Excavation in the new millenium*. Hadrian Books. pp.265–268.

Hacking, I. (1975). *The emergence of probability: a philosophical study of early ideas about probability, induction and statistical inference*. Cambridge University Press.

Hacking, I. (1983). *Representing and intervening: introductory topics in the philosophy of natural science*. Cambridge University Press.

Hammer, F. (2000). The structuring of records: different systems in use from excavation through to publication. In: S. Roskams (Ed). *Interpreting Stratigraphy: Site evaluation, recording procedures and stratigraphic analysis.* Archaeopress. pp.143–149.

Hanson, N. R. (1958). *Patterns of discovery: an inquiry into the conceptual foundations of science*. Cambridge University Press.

Hanson, W. S. and Rahtz, P. A. (1988). Video recording on excavations. *Antiquity*, 62 (234), pp.106–111. [Online]. Available at: doi:10.1017/S0003598X00073567.

Harris, E. (1989). *Principles of archaeological stratigraphy*. 2nd ed. London : Academic Press.

Harris, E., Brown, M. and Brown, G. (1993). *Practices of archaeological stratigraphy*. London : Academic Press.

Hawkes, C. (1954). Archaeological Theory and Method: Some Suggestions Theory Archeological from the Old World. *American Anthropologist*, 56 (2), pp.155–168.

Heggie, V. (2014). Why Isn’t Exploration a Science? *Isis: A Journal of the History of Science*, 105 (2), pp.318–334.

Historic England. (2015). *Management of Research Projects in the Historic Environment*.

Hodder, I. (1986). *Reading the past: current approaches to interpretation in archaeology*. Cambridge University Press.

Hodder, I. (1991). Interpretive archaeology and its role. *American Antiquity*, 56 (1), pp.7–18.

Hodder, I. (1997). 'Always momentary, fluid and flexible': Towards a reflexive excavation methodology. *Antiquity*, 71 (273), pp.691–700. [Online]. Available at: doi:10.1017/S0003598X00085410.

Hodder, I. (1999). *The archaeological process: an introduction*. Blackwell.

Hodder, I. (2000a). Developing a reflexive method in archaeology. In: Hodder, I. (Ed). *Towards reflexive method in archaeology: the example at Çatalhöyük*. Cambridge : McDonald Institute for Archaeological Research. pp.3–18.

Hodder, I. (2000b). The interpretation of documents and material culture. In: Denzin, N. K. and Lincoln, Y. S. (Eds). *Handbook of qualitative research*. Sage. pp.703–716.

Hodder, I. (2005). Reflexive method. In: H. Maschner (Ed). *Handbook of archaeological methods. Volume 1*. Altamira Press. pp.643–669.

Hodder, I. (2006). *Excavating Çatalhöyük: South, North and Kopal area reports from the 1995-1999 seasons*. I. Hodder (Ed). Cambridge : McDonald Institute for Archaeological Research.

Hodder, I. (Ed). (2014). *Çatalhöyük excavations: the 2000-2008 seasons*. British Institute at Ankara, Cotsen Institute of Archaeology at UCLA.

Hodder, I. and Farid, S. (2014). Questions, history of work and summary of results. In: I. Hodder (Ed). *Çatalhöyük excavations : the 2000-2008 seasons*. British Institute at Ankara. pp.1–34.

Hodder, I. and Hutson, S. (2003). *Reading the Past: Current Approaches to Interpretation in Archaeology*. 3rd ed. Cambridge University Press.

Hon, G. (1995). Going Wrong: To Make a Mistake , to Fall into an Error. *Review of Metaphysics*, 49 (1), pp.3–20.

Hook, B. et al. (2016). *National infrastucture development and historic enviroment skills and capacity 2015-33: An assesment*. Historic England,.

Hope-Taylor, B. (1966). Archaeological Draughtsmanship: Principles and Practice. Part II: Ends and Means. 2015/01/02. *Antiquity*, 40 (158), Cambridge University Press., pp.107–113. [Online]. Available at: doi:DOI: 10.1017/S0003598X00032270.

Hope-Taylor, B. (1977). *Yeavering: an Anglo-British centre of early Northumbria*. London : English Heritage.

Jennings, D. (2011). Great excavations , developer funding and the future. In: J. Schofield (Ed). *Great excavations: Shaping the archaeological profession*. Oxbow Books. pp.279–288.

Johnson, M. (2010). *Archaeological theory: an introduction*. 2nd ed. Wiley-Blackwell.

Johnson, M. H. (2006). Archaeology and theoretical culture. *Archaeological Dialogues*, 13 (2), pp.167–182. [Online]. Available at: doi:10.1017/S1380203806292080.

Jones, A. M. (2013). 'It started with things'. An interview with Professor Richard Bradley. *Archaeological Dialogues*, 20 (2), pp.175–185. [Online]. Available at: doi:10.1017/S1380203813000226.

Jones, B. (1984). *Past imperfect : the story of rescue archaeology*. London : Heinemann.

Jones, M. (2000). The stratigraphic examination of standing buildings. In: S. Roskams (Ed). *Interpreting Stratigraphy: Site evaluation, recording procedures and stratigraphic analysis.* Archaeopress. pp.113–123.

Joyce, R. A. and Pollard, J. (2010). Archaeological Assemblages and Practices of Deposition. In: D. Hicks and Beaudry, M. (Eds). *The Oxford Handbook of Material Culture Studies*. Oxford University Press. pp.292–309. [Online]. Available at: doi:10.1093/oxfordhb/9780199218714.013.0012.

Kenyon, K. (1974). *Digging up Jerusalem*. New York : Praeger.

Knight, M. et al. (2019). The Must Farm pile-dwelling settlement. 2019/06/12. *Antiquity*, 93 (369), Cambridge University Press., pp.645–663. [Online]. Available at: doi:DOI: 10.15184/aqy.2019.38.

Kosso, P. (2001). *Knowing the Past: Philosophical Issues of History and Archaeology*. Humanity Books.

Kristiansen, K. (2009). Contract archaeology in Europe: An experiment in diversity. *World Archaeology*, 41 (4), pp.641–648. [Online]. Available at: doi:10.1080/00438240903371486.

Krotscheck, U. (2003). Space 100. In: *Catalhoyuk 2003 Archive Report*. Catalhoyuk Research Project. pp.23–26. [Online]. Available at: http://www.catalhoyuk.com/research/archive_reports.

Krotscheck, U. (2004). Building 49. In: *Catalhoyuk 2004 Archive Report*. Catalhoyuk Research Project. [Online]. Available at: http://www.catalhoyuk.com/archive_reports/2004/.

Ladyman, J. (2002). *Understanding philosophy of science*. London : Routledge.

Lakin, D. (1995). *Northgate house 20-28 Moorgate, EC2. City of London: An archaeological evaluation*.

Leighton, M. (2015). Excavation methodologies and labour as epistemic concerns in the practice of archaeology. Comparing examples from British and Andean archaeology. *Archaeological Dialogues*, 22 (1), pp.65–88. [Online]. Available at: doi:10.1017/S1380203815000100.

Levi, G. (2001). On microhistory. In: P. Burke (Ed). *New perspective on historical writting*. Polity press. pp.97–119.

Lewis, J. (Editor). (2006). *Landscape evolution in the Middle Thames Valley: Heathrow Terminal 5 excavations*. Oxford and Salisbury : Framework Archaeology.

List, C. (2011). Group Knowledge and Group Rationality: A Judgment Aggregation Perspective. In: Goldman, A. and D. Whitcomb (Eds). *Social Epistemology: Essential Readings*. 2012/01/03. Oxford University Press. pp.221–241. [Online]. Available at: doi:DOI: 10.3366/epi.2005.2.1.25.

Livingstone, D. (2003). *Putting science in its place: geographies of scientific knowledge*. University of Chicago Press.

Lucas, G. (2001). *Critical approaches to fieldwork: contemporary and historical archaeological practice*. London : Routledge.

Lucas, G. (2009). Modes of recording. In: C. Evans (Ed). *Fengate revisited: further Fen-edge excavations, bronze age fieldsystems and settlement and the Wyman Abbott/Leeds Archives*. Cambridge Archaeological Unit. p.232.

Lucas, G. (2012a). Fieldwork and Collecting. *The Oxford Handbook of Material Culture Studies*, (March 2017), pp.6–8. [Online]. Available at: doi:10.1093/oxfordhb/9780199218714.013.0009.

Lucas, G. (2012b). *Understanding the archaeological record*. Cambridge University Press.

Lucas, G. (2014). Interpretation in Archaeological Theory. In: C. Smith (Ed). *Encyclopedia of Global Archaeology*. New York : Springer New York. pp.4013–4022. [Online]. Available at: doi:10.1007/978-1-4419-0465-2_290.

Lucas, G. (2015). Evidence of what? On the possibilities of archaeological interpretation. In: R. Chapman and A. Wylie (Eds). *Material Evidence Learning from Archaeological Practice*. Routledge. pp.311–323.

Lucas, G. (2019). *Writing the past : knowledge and literary production in archaeology*. Routledge.

Lukas, D., Engel, C. and Mazzucato, C. (2018). Towards a Living Archive: Making Multi Layered Research Data and Knowledge Generation Transparent. *Journal of Field Archaeology*, 43 (sup1), pp.S19–S30. [Online]. Available at: doi:10.1080/00934690.2018.1516110.

May, S. (2012). Commentary. In: H. Cobb et al. (Eds). *Reconsidering Archaeological Fieldwork: Exploring On-Site Relationships between Theory and Practice*. pp.167–179. [Online]. Available at: doi:10.1007/978-1-4614-2338-6.

McAdam, E. (1992). Discussion Session. Feedback Mechanisms from Post-Excavation to Excavation. In: *Interpreting Stratigraphy 3. A Review of the Art*. p.7.

McAnany, P. A. and Hodder, I. (2009). Thinking about stratigraphic sequence in social terms. *Archaeological Dialogues*, 16 (1), pp.1–22. [Online]. Available at: doi:10.1017/S1380203809002748.

Mellaart, J. (1967). *Çatal Hüyük: a neolithic town in Anatolia*. London : Thames & Hudson.

Merleau-Ponty, M. (2011). *Phenomenology of perception*. London : Routledge.

Mickel, A. (2015). Reasons for Redundancy in Reflexivity: The Role of Diaries in Archaeological Epistemology. *Journal of Field Archaeology*, 40 (3), pp.300–309. [Online]. Available at: doi:10.1179/2042458214Y.0000000002.

Milner, N., Conneller, C. and Taylor, B. (2018). *Star Carr Volume 1*. Milner, N., Conneller, C. and Taylor, B. (Eds). York : White Rose University Press. [Online]. Available at: doi:10.22599/book1 SE - 407.

Morgan, C. and Wright, H. (2018). Pencils and Pixels: Drawing and Digital Media in Archaeological Field Recording. *Journal of Field Archaeology*, 43 (2), pp.136–151. [Online]. Available at: doi:10.1080/00934690.2018.1428488.

Moser, S. (2014). Making Expert Knowledge through the Image: Connections between Antiquarian and Early Modern Scientific Illustration. *Isis*, 105 (1), [The University of Chicago

Press, The History of Science Society]., pp.58–99. [Online]. Available at: doi:10.1086/675551.

Mulligan, L. (1996). Self-Scrutiny and the Study of Nature: Robert Hooke's Diary as Natural History. *Journal of British Studies*, 35 (3), pp.311–342.

Museum of London. (1994). *Archaeological Site Manual. Museum of London Archaeological Service*. Third. London .

Museum of London. (2000). *Northgate house and Kent house*. London .

Olsen, B. et al. (2012). *Archaeology: the Discipline of Things*. Berkerley : University of California Press.

Ottaway, P. (1992). *Archaeology in British towns : from the Emperor Claudius to the Black Death*. London : Routledge.

Ottaway, P. (2011). *Archaeology in the environs of Roman York : excavations 1976-2005*. York Archaeological Trust.

Pavel, C. (2010a). *Describing and interpreting the past: European and American approaches to the written record of the excavation*. Bucureşti : Editura Universităţii din Bucureşti.

Pavel, C. (2010b). Fişele stratigrafice („context sheets") în săpătura arheologică. Perspective critice. *SCIVA*, 61, pp.251–279.

Petrie, F. (1904). *Methods and aims in archaeology*. Cambridge University Press.

Piggott, S. (1965). Archaeological Draughtsmanship: Principles and Practice Part I: Principles and Retrospect. 2015/01/02. *Antiquity*, 39 (155), Cambridge University Press., pp.165–176. [Online]. Available at: doi:DOI: 10.1017/S0003598X00031823.

Popper, K. R. (1972). *Conjectures and refutations: the growth of scientific knowledge*. 4th ed. London : Routledge and Kegan Paul.

Quine, W. V. (1969). *Ontological relativity and other essays*. New York : Columbia University Press.

Radder, H. (2006). *The world observed, the world conceived*. University of Pittsburgh Press.

Rahtz, P. (1974a). Volunteers. In: P. Rahtz (Ed). *Rescue archaeology*. Penguin Books. pp.274–279.

Rahtz, P. A. (1974b). *Rescue archaeology*. Harmondsworth : Penguin.

Reynolds, N. and Barber, J. (1984). Analytical excavation. 2015/01/02. *Antiquity*, 58 (223), Cambridge University Press., pp.95–102. [Online]. Available at: doi:DOI: 10.1017/S0003598X00051504.

Roosevelt, C. H. et al. (2015). Excavation is Destruction Digitization: Advances in Archaeological Practice. *Journal of Field Archaeology*, 40 (3), pp.325–346. [Online]. Available at: doi:10.1179/2042458215Y.0000000004.

Roskams, S. (2000). *Interpreting stratigraphy: site evaluation, recording procedures and stratigraphic analysis: papers presented to the Interpreting Stratigraphy Conferences, 1993-*

1997. Oxford : Archaeopress.

Roskams, S. (2001). *Excavation*. Cambridge University Press.

Sapirstein, P. and Murray, S. (2017). Establishing Best Practices for Photogrammetric Recording During Archaeological Fieldwork. *Journal of Field Archaeology*, 42 (4), pp.337–350. [Online]. Available at: doi:10.1080/00934690.2017.1338513.

Schiffer, M. B. (1972). Archaeological Context and Systemic Context. *American Antiquity*, 37 (2), pp.156–165. [Online]. Available at: doi:10.2307/278203.

Schiffer, M. B. (1987). *Formation processes of the archaeological record*. 1st ed. Albuquerque : University of New Mexico Press.

Schlanger, N. and Aitchison, K. (Eds). (2010). *Archaeology and the Global Economic Crisis: Multiple Impacts, Possible Solutions*. Culture Lab Editions.

Seeley, F. and Drummond-Murray, J. (2005). *Roman pottery production in the Walbrook valley : excavations at 20-28 Moorgate, City of London, 1998-2000*. London : Museum of London Archaeology Service.

Shanks, M. and Hodder, I. (1994). Processual, postprocessual and interpretive archaeologies. In: Hodder, I. (Ed). *Interpreting archaeology: finding meaning in the past*. London : Routledge. pp.3–29.

Shanks, M. and McGuire, R. (1996). The craft of archaeology. *American Antiquity*, 61 (1), pp.75–88. [Online]. Available at: doi:10.1161/01.STR.32.1.139.

Shanks, M. and Tilley, C. (1992). *Re-constructing archaeology: theory and practice*. 2nd ed. London : Routledge.

Shapin, S. (1989). The invisible technician. *American Scientist*, 77, pp.554–563.

Shapin, S. (1994). *A social history of truth: civility and science in seventeenth-century England*. Chicago : University of Chicago Press.

Silverman, D. (2011). *Interpreting qualitative data: a guide to the principles of qualitative research*. 4th ed. London : Sage.

Spall, C. (2006). *Research excavation report. Tarbat discovery program. Data structure report March 2006*. City of York . [Online]. Available at: doi:https://doi.org/10.5284/1031216.

Spall, C. (2007). *Research excavation report. Tarbat discovery programme. Data structure report March 2007*. City of York . [Online]. Available at: doi:https://doi.org/10.5284/1031216.

Spall, C., Garner-Lahire, J. and Carver, M. (2012). *Research excavation. Sector 2 and 3. Data structure report*. York . [Online]. Available at: doi:https://doi.org/10.5284/1031216.

Spence, C. (1993). Recording the archaeology of London: the development an implementation of the DUA recording system. In: E. Harris, Brown, R. and G. Brown (Eds). *Practices of archaeological stratigraphy*. pp.23–46.

Steane, K. et al. (1992). *Interpretation of stratigraphy: a review of the art. Proceedings of a conference held on 18th June, 1992, at the City of Lincoln Archaeology Unit*. City of Lincoln

Archaeology Unit.

Stocking, G. W. (1983). The etnographer's magic: Fieldwork in British anthropology from Tylor to Malinowski. In: *Observers observed : essays on ethnographic fieldwork*. University of Wisconsin Press. pp.70–120.

Tassie, G. (2015). Single-Context Recording in the context of archaeological fieldwork in Egypt, part one. In: Hassan, F. A. et al. (Eds). *The Management of Egypt's Cultural Heritage*. London : ECHO and Golden House Publications. pp.175-184.

Taylor, J. et al. (2018). The Rise of the Machine': the impact of digital tablet recording in the field at Çatalhöyük. *Internet Archaeology*, (47). [Online]. Available at: https://doi.org/10.11141/ia.47.1.

Thomas, J. (1993). The hermeneutics of megalithic space. In: C. Tilley (Ed). *Interpretative Archaeology*. Berg.

Thomas, J. (2004). The Great Dark Book: Archaeology , Experience , and Interpretation. In: J. Bintliff (Ed). *A companion to Archaeology*. Blackwell. pp.21–36.

Thomas, J. (2013). *The birth of Neolithic Britain: an interpretive account*. Oxford University Press.

Thorpe, R. (2012). Often fun, usually messy: fieldwork, recording and higher order of things. In: H. Cobb et al. (Eds). *Reconsidering Archaeological Fieldwork: Exploring On-Site Relationships between Theory and Practice*. Springer. pp.31–52. [Online]. Available at: doi:10.1007/978-1-4614-2338-6.

Tilley, C. (1989). Excavation as theatre. *Antiquity*, 63 (239), pp.275–280. [Online]. Available at: doi:10.1017/S0003598X00075992.

Tilley, C. (1993). Introduction : Interpretation and a Poetics of the Past. In: C. Tilley (Ed). *Interpretative Archaeology*. London : Bloomsbury Academic. pp.1–28. [Online]. Available at: http://www.bloomsburycollections.com/book///.

Tilley, C. (2009). *Interpreting landscapes: geologies, topographies, identities*. Walnut Creek, Calif. : Left Coast Press.

Tomášková, S. (2006). On being heard. Theory as an archaeological practice. *Archaeological Dialogues*, 13 (2), pp.163–167. [Online]. Available at: doi:DOI: 10.1017/S1380203806282084.

Trigger, B. (2006). *A history of archaeological thought*. 2nd ed. Cambridge University Press.

Trigham, R. and Stevanovic, M. (2000). Different excavation styles create different windows into Catalhoyuk. In: I. Hodder (Ed). *Towards reflexive method in archaeology: the example at Çatalhöyük*. McDonald Institute for Archaeological Research. pp.111–118.

Turner, D. (2007). *Making Prehistory: Historical Science and the Scientific Realism Debate*. Cambridge University Press. [Online]. Available at: doi:DOI: 10.1017/CBO9780511487385.

Watson, S. (2019). Whither archaeologists? Continuing challenges to field practice. 2019/09/20. *Antiquity*, 93 (372), Cambridge University Press., pp.1643–1652. [Online]. Available at: doi:DOI: 10.15184/aqy.2019.141.

Weirich, P. (2010). *Collective Rationality: Equilibrium in Cooperative Games*. New York : Oxford University Press. [Online]. Available at: doi:10.1093/acprof:oso/9780195388381.001.0001.

Wheeler, M. (1943). *Maiden Castle, Dorset.* London, Society of Antiquaries .

Woodward, J. (1989). Data and phenomena. *Synthese*, 79 (3), Springer., pp.393–472.

Woodward, J. (2000). Data , Phenomena , and Reliability. *Philosophy of science*, 67, pp.163–179.

Wylie, A. (2002). *Thinking from things: essays in the philosophy of archaeology*. Berkeley : University of California Press.

Yin, R. K. (2014). *Case study research: design and methods*. 5th ed. SAGE.

Zemon-Davis, N. (1983). *The return of Martin Guerre*. Cambridge, Mass. : Harvard University Press.

Zubrow, E., Audouze, F. and Enloe, J. G. (Eds). (2010). *The Magdalenian household: unraveling domesticity*. Albany : State University of New York Press.