A thesis submitted for the degree of Doctor of Philosophy in Archaeology

The skill spectrum

Methods for identifying variable skill-levels amongst handaxe-making hominins in the Middle Pleistocene

Laurie Hutchence

Wolfson College
University of Oxford
Hilary Term 2022

Some parts of this thesis cover work previously published in scholarly journals:

Hutchence, L. & Debackere, S. (2018). An evaluation of behaviours considered indicative of skill in handaxe manufacture. *Lithics: the journal of the Lithic Studies Society* 39: 36–51.

Hutchence, L., & Scott, C. (2021). Is Acheulean Handaxe Shape the Result of Imposed 'Mental Templates' or Emergent in Manufacture? Dissolving the Dichotomy through Exploring 'Communities of Practice' at Boxgrove, UK. *Cambridge Archaeological Journal, 31*(4), 675-686. doi:10.1017/S0959774321000251

Abstract

Handaxes are bifacial stone cutting tools that were made by hominins between circa 1.75 and 0.2 million years ago. They represent a significant advancement in stone-tool technology and a pivotal development in hominin cognition and behaviour. Experimental research suggests learning to make handaxes demands considerable repeated practice over time and social support mechanisms that encourage skill development. Considerable debate exists regarding whether guidance for learners in the Acheulean included verbal and/or gestural instruction, or neophytes learned solely through imitation/emulation. Whatever the mode of transmission, consistency in intra-assemblage biface morphology and manufacturing techniques implies normative standards for acceptable performance were maintained, even with resharpening, at sites like Gouzeaucourt (the focus of this thesis). Assemblages like this, where knappers shared common goals, permit the identification of the varying ability of individuals to adhere to such objectives. Additionally, because mastery of handaxe manufacture requires extensive training, it follows that a spectrum of skill should be visible in the archaeological record through analysing the bifaces that remain. Nonetheless, few studies have applied experimental outcomes directly to interpreting archaeological bifaces and the differences in skill-level they exhibit. It is this lacuna in Acheulean research that is addressed here. My results show that the quantitative and qualitative methodology I employ allows accurate differentiation between skilled and unskilled experimental knappers and, as such, helps to categorise archaeological bifaces according to the perceived skill-level of their maker. I have thus presented an original approach to collecting data from bifaces regarding previously unknown aspects of the hominins that made them. This information can be used to calculate the proportion of skill-levels contributing to an assemblage, thus providing more inventive ways of interpreting Palaeolithic sites. Consequently, I thereby

anticipate that this thesis represents a prospective starting point for further experimental investigations into differentiating between skill-levels in the Acheulean.

Preface and acknowledgements

Stone knapping requires repeated practice to acquire expertise. A knapper's aptitude is associated with the amount of experience they have accumulated. As each knapper continues to practice, they will recognise a series of progressions in their own execution of certain behaviours that characterise each stage of their skill development, along a spectrum from beginner to expert. For many skilled activities however, the upshots of performance are momentary and for the assessor differences in skill can only be judged via observation of fleeting bodily movements. Alternatively, skilled crafts provide the opportunity to recover physical manifestations of performance indicative of an agent's aptitude. This is advantageous for archaeologists who are not privy to the action of individuals in the past. Artefacts thus possess the potential to demonstrate differences in the skill-levels of our hominin ancestors.

In the Acheulean, discussion of skill has largely been related to developments in technological advancements in handaxe manufacture. Occasionally, particularly well made, or poorly made, bifaces are remarked upon, but there are a few systematic approaches to differentiating bifaces by skill-level in the archaeological record. This is despite the fact there is a wealth of experimental research aimed at doing just that among modern knappers. Indeed, experimental analysis has highlighted key indicators of skilled/unskilled performance. However, the only methodology aimed at capturing this important aspect of hominin life is so far provided by Leroyer (2016), who does not utilise experimental data in his assessments. Therefore, I advance a methodology in this thesis that employs the central findings of my own experimental investigation to help me to discriminate between archaeological bifaces by the skill-level category that can be credited to their maker.

My interest in this aspect of hominin life was sparked by reading Jacques Pelegrin and conducting my first knapping experiments as an undergraduate. Since then, I have been able to pursue my interest to this point with the help of many great people, whose energy, insight, and unerring patience helped replenish my resolve at crucial points along the way.

I would like to express my thanks to Dr Agnes Lamotte, Université de Lille, who supervised my time analysing bifaces from Gouzeaucourt at the Maison de l'Archeologie in Arras, France. I have continually relied on her knowledge and offer of support throughout my DPhil. In addition, I am grateful to Mme Armelle Masse, Head of the Soil Archives Service, Pas-de-Calais CCE, for her help in accessing the material.

Many thanks are due to Chris Scott and Prof. Larry Barham also at the University of Liverpool, who kindly provided me with a space to work with their archive of experimental knapping material. I am also greatly appreciative of everyone who participated in my experiments, as well as those who completed my ethnographic questionnaire. Additionally, I am indebted to Antonis Iliopoulous for always being available to discuss theoretical questions, and for helping me formulate crucial ideas. Moreover, many thanks go to my supervisor, Prof. Nick Barton, who has provided critical help and was instrumental in improving my clarity of expression and presentation.

This work could not have been achieved without the funding provided by an AHRC DTP Studentship award and the additional funding for fieldwork provided by AHRC and Wolfson College Travel Grants.

Lastly, I am beholden to the legion of friends and family who have taken time out of their busy schedules to support me. It is to them that this thesis is dedicated.

Table of Contents

Abstract ..... ***ii***

Preface and acknowledgements ..... ***iv***

List of Figures ..... ***x***

List of Tables ..... ***xii***

List of Appendix Figures ..... ***xiv***

List of Appendix Tables ..... ***xiv***

1. Chapter 1 – Introduction ***1***

1.1 Overview **1**

1.2 Introducing Gouzeaucourt **3**

1.3 Approach **6**

1.4 Thesis Structure **8**

2. Chapter 2 – The Acheulean ***11***

2.1 First principles: who, what, where, when? **11**

2.2 What is a biface? **12**

2.3 Evolutionary Significance of bifacial technology **18**

2.3.1 Cognition 20

2.3.2 Social Learning 23

2.3.3 Utilitarian vs. non-utilitarian significance 28

2.3 Summary **34**

3. Chapter 3 – Handaxe morphological variation: a history of debate ***37***

3.1 Introduction **37**

3.2 Acheulean conservatism **38**

3.3 Principal Approaches to Morphological Variability **48**

3.3.1 Culture Historical Approach 48

3.3.2 Utilitarian Approach 50

3.3.3 Holistic Approach 58

3.3.4 Summary 61

3.4 Relevance of skill-level to debate **63**

4. Chapter 4 – Theoretical framework ***67***

4.1 Definitions of and approaches to understanding skill in the Acheulean **67**

4.2 Approaches to understanding skill **69**

4.2.1 Knowledge and know-how 69

4.2.2 Traditional internalist versus Material Engagement Theory (MET) 71

4.2.3 A Neurophysiological perspective 78

4.2.4 Plateaus, Dips, and Leaps (PDLs) 84

4.3 Ethnographic case studies **90**

4.3.1 Roux and Corbetta, 1989 90

4.3.2. Keller and Keller, 1996 92

4.3.3. Biryukova et al., 2015 95

4.4 Experimental archaeological case studies **97**

4.5 The relevance of the 'social' in discussions of skill 101

4.6 The skill spectrum: the theoretical reality 103

5. Chapter 5 – Research Rationale and Method of Approach 105

5.1 Introduction 105

5.2 Experimental Investigation 106

5.3 Archaeological Investigation 108

5.4. Gouzeaucourt Level G (GznivG) 111
5.4.1. Contextualising Gouzeaucourt 111
5.4.2. Gouzeaucourt overview 114
5.4.3. Challenging a utilitarian perspective of Gouzeaucourt 118

5.5 Morphological homogeneity at Gouzeaucourt 124
5.5.1. Rationale 125
5.5.2. Approach 125
5.5.3. Evaluation 127

5.6 Summary 133

6. Chapter 6 - Methodology 135

6.1 Introduction 135

6.2 Methodological approach 136

6.3 Materials 137
6.3.1 Experimental assemblage 137
6.3.2 Archaeological assemblage 142
6.3.3 Questionnaire 143

6.4 Data collection 144
6.4.1 Quantitative data 145
6.4.2 Qualitative data 152
6.4.3 Diacritical data 154
6.4.4 Questionnaire data 155

6.5 Analyses 155
6.5.1 Experimental assemblage 156
6.5.2 Archaeological assemblage 163

6.6 Summary 170

7. Chapter 7 – Results 172

7.1 Introduction 172

7.2 Experimental Investigation 172
7.2.1 Quantitative analysis 173
7.2.2 Qualitative Analysis 237

7.3 Blind Test 256

7.4 Archaeological Investigation 261
7.4.1 Quantitative Analysis 261
7.4.2 Qualitative classification and comparison with quantitative results 282

7.5 Summary 305
7.5.1 Main experimental investigation 307
7.5.2 Archaeological Hypotheses 308

8. Chapter 8 – Discussion and Interpretation 311

8.1 Introduction 311

8.2 Key findings and limitations **313**

8.3 Further testing of my methodology – Boxgrove, UK **320**

8.3.1 Boxgrove LDA 321

8.3.2 Post-LDA univariate statistics 325

8.3.3 Qualitative analysis 326

8.4 Main outcomes of my research and their relation to areas of future study **334**

8.5 Assessing site function **338**

8.6 Conclusion **341**

Bibliography ***343***

Appendix. * ***358***

List of Figures

Figure 1. 1. Location of Gouzeaucourt site 5
Figure 2. 1. Planform and profile views of Early Acheulean handaxe……13
Figure 2. 2. Planform and Profile views of a Later Acheulean handaxe 13
Figure 2. 3. Planform and Profile view of a Cleaver 15
Figure 2. 4. Range of handaxe shapes (tip shapes) as defined by Roe (1968) 16
Figure 2. 5. Distal (tip) ends encountered at Boxgrove 19
Figure 2. 6. Planform outlines of a randomized sample of handaxes from Boxgrove, Unit 4c. 32
Figure 2. 7. Homogeneity of planform shape at Boxgrove. 33
Figure 3. 1. Emery's (2010) appraisal of Ashton and White's (2003) critique of McPherron's (1994) resharpening hypothesis…….53
Figure 3. 2. All factors that influence handaxe shape. 59
Figure 3. 3. Leroyer's (2016) skill-spectrum at Boxgrove. 65
Figure 5. 1. Example of secondary reworking on Biface 8015(9)……117
Figure 5. 2. Consistency of biface shape at GZnivG 118
Figure 5. 3. Grouped Scatter graph correlation of Length vs L1/L at GZnivG sorted by biface shape. 122
Figure 5. 4. Illustration of XY coordinate data captured via AScheck method 126
Figure 5. 5. Principal component scatter graph showing shape variance at Gouzeaucourt (G, H and R) Boxgrove and Broom Pits assemblages 131
Figure 5. 6. Outlines of all bifaces from samples from GZnivG (a), Boxgrove (b), Broom Pits (C), and GZnivR (d). 132
Figure 5. 7. Outlines of longest and shortest 10% of biface in samples from GZnivG (a), Boxgrove (b), Broom Pits (C), and GZnivR (d). 132
Figure 6. 1. Blind Test example biface made by John Lord. ……..142
Figure 6. 2. AScheck process of calculating symmetry 149
Figure 6. 3. Illustration of the measurements needed to calculate Edge straightness. 150
Figure 6. 4. Flake terminations. 153
Figure 6. 5. Diacritical analysis as per Lamotte's (2012, 2011) method. 154
Figure 7. 1. Box and whisker plot showing range of experimental biface thickness by skill-level……...174
Figure 7. 2. Box and whisker plot showing range of experimental biface L1/L ratios by skill-level. 178
Figure 7. 3. Boxplot showing range of Flake Scar number scores by skill-level. 181
Figure 7. 4. Scatter plot of LDA 1 jackknifed experimental biface scores grouped by reclassified skill-level (including Intermediate Drafts). 203
Figure 7. 5. Scatter plot of LDA 1 jackknifed experimental biface scores grouped by reclassified skill-level (excluding Intermediate Drafts). 205
Figure 7. 6. Scatter plot of LDA 2 jackknifed experimental biface scores grouped by reclassified skill-level (including Intermediate Drafts). 207
Figure 7. 7. Scatter plot of LDA 2 jackknifed experimental biface scores grouped by reclassified skill-level (excluding Intermediate Drafts). 209
Figure 7. 8. Scatter plot of LDA 3 jackknifed experimental biface scores grouped by reclassified skill-level (including Intermediate Drafts). 212
Figure 7. 9. Scatter plot of LDA 3 jackknifed experimental biface scores grouped by reclassified skill-level (excluding Intermediate Drafts). 214
Figure 7. 10. Scatter plot of LDA 4 jackknifed experimental biface scores grouped by reclassified skill-level (including Intermediate Drafts). 217

Figure 7. 11. Scatter plot of LDA 4 jackknifed experimental biface scores grouped by reclassified skill-level (excluding Intermediate Drafts). 218
Figure 7. 12. Scatter plot of LDA 1 (ovates only) jackknifed experimental biface scores grouped by reclassified skill-level (including Intermediate Drafts). 226
Figure 7. 13. Scatter plot of Experimental knapper 7's jackknifed scores grouped by reclassified session ID. 230
Figure 7. 14. Biface CS_LIV1.1 in Planform and Profile views. 241
Figure 7. 15. Biface CS_LIV1.4 in Planform and Profile views. 242
Figure 7. 16. Biface CS_LIV1.6 in Planform and Profile views. 243
Figure 7. 17. Biface KL_LIV.20 in Planform and Profile views. 245
Figure 7. 18. Biface CS_LIV3.12 in Planform and Profile views. 246
Figure 7. 19. Biface CS_LIV2.13 in Planform and Profile views. 248
Figure 7. 20. Biface WM_JL.1 in Planform and Profile views. 250
Figure 7. 21. Biface LH_KL.2 in Planform and Profile views. 253
Figure 7. 22. Biface LH_PSTKL.1 in Planform and Profile views. 253
Figure 7. 23. PCA scatter plot of Blind test bifaces and the example biface. 259
Figure 7. 24. Scatter plot of LDA re-classification of GZnivG bifaces. 275
Figure 8. 1. LDA scatter plot of Boxgrove bifaces by initial classification…… ...323

List of Tables

Table 5. 1. Discriminant Analysis of biface shape variance at Gouzeaucourt (Levels G, H and R), Boxgrove, and Broom Pits, showing percentage of correct classification by assemblage. 130
Table 6. 1. Example LDA Axis loading scores by criterion. 165
Table 6. 2. Example LDA Axis score by biface. 165
Table 6. 3. Diacritical assessment labels. 170
Table 7. 1. LDA 1: Skill criteria loadings (including Intermediate Drafts). ……..202
Table 7. 2. LDA 1: Confusion Matrix (including Intermediate Drafts). 203
Table 7. 3. LDA 2: Skill criteria loadings (excluding Intermediate Drafts). 204
Table 7. 4. LDA 1: Confusion Matrix (excluding Intermediate Drafts). 205
Table 7. 5. LDA 2: Skill criteria loadings (including Intermediate Drafts). 206
Table 7. 6. LDA 2: Confusion Matrix (including Intermediate Drafts). 207
Table 7. 7. LDA 2: Skill criteria loadings (excluding Intermediate Drafts). 208
Table 7. 8. LDA 2: Confusion Matrix (excluding Intermediate Drafts). 209
Table 7. 9. LDA 3: Skill criteria loadings (including Intermediate Drafts). 211
Table 7. 10. LDA 3: Confusion Matrix (including Intermediate Drafts). 211
Table 7. 11. LDA 3: Skill criteria loadings (excluding Intermediate Drafts). 212
Table 7. 12. LDA 3: Confusion Matrix (excluding Intermediate Drafts). 213
Table 7. 13. LDA 4: Skill criteria loadings (including Intermediate Drafts). 215
Table 7. 14. Table 7.14. LDA 4: Confusion Matrix (including Intermediate Drafts). 216
Table 7. 15. LDA 4: Skill criteria loadings (excluding Intermediate Drafts). 217
Table 7. 16. LDA 4: Confusion Matrix (excluding Intermediate Drafts). 218
Table 7. 17. Summary of LDA accuracy (including and excluding Intermediate Drafts). 219
Table 7. 18. Summary of LDA accuracy. 222
Table 7. 19. LDA 1 ovate only: Skill criteria loadings (including Intermediate Drafts). 224
Table 7. 20. LDA 1 ovate only: Confusion matrix (including Intermediate Drafts). .. 225
Table 7. 21. LDA 1 skill criteria loadings for Experimental Knapper 7's assemblage. 229
Table 7. 22. Confusion matrix of Experiment knapper 7's bifaces. 230
Table 7. 23. LDA 1 skill criteria loadings for Experimental Knapper 4's assemblage (including Intermediate Drafts). 233
Table 7. 24. Confusion matrix of Experiment knapper 4's bifaces (including Intermediate Drafts). 234
Table 7. 25. LDA 1 skill criteria loadings by Axis for Experimental Knapper 4's assemblage (excluding Intermediate Drafts). 235
Table 7. 26. Confusion matrix of Experiment knapper 4's bifaces (excluding Intermediate Drafts). 236
Table 7. 27. Comparison of original and jackknifed classifications for Experimental ovate bifaces. 240
Table 7. 28. LDA 1 skill criteria loadings for Blind Test biface assemblage. 258
Table 7. 29. PCA loadings of Blind Test bifaces. 258
Table 7. 30. Skill-level classification given to Blind Test bifaces. 260
Table 7. 31. Comparison of Ratio of refinement univariate statistics between Experimental and GZnivG assemblages. 264
Table 7. 32. Comparison of Th/L univariate statistics between Experimental and GZnivG assemblages. 264

Table 7. 33. Comparison of Planform Asymmetry univariate statistics between Experimental and GZnivG assemblages. 265
Table 7. 34. Comparison of Profile Asymmetry univariate statistics between Experimental and GZnivG assemblages. 265
Table 7. 35. Comparison of Edge straightness univariate statistics between Experimental and GZnivG assemblages. 266
Table 7. 36. Comparison of L1/L univariate statistics between Experimental and GZnivG assemblages. 267
Table 7. 37. Means and ranges of experimental skill-levels by normalised Axes scores. 270
Table 7. 38. Skill-level thresholds applied to GZnivG based on normalised experimental data. 270
Table 7. 39. Initial quantitative skill-level assessment of GZnivG bifaces. 271
Table 7. 40. LDA 1 normalised GZnivG axis loading scores. 273
Table 7. 41. Confusion Matrix by normalised skill-level classification of GZnivG bifaces. 273
Table 7. 42. GZnivG: comparison of quantitative and qualitative skill-level assessments. 302
Table 7. 43. Diacritical assessment by qualitative skill-level classification. 303
Table 8. 1. Initial quantitative skill-level classification of Boxgrove sample. 321
Table 8. 2. Loading scores by axis of normalised Boxgrove axes. 322
Table 8. 3. Confusion matrix of Boxgrove sample bifaces. 323
Table 8. 4. Qualitative corroboration of quantitative skill-level assessments of Boxgrove bifaces. 329

List of Appendix Figures

Appendix Figure 1. Example LDA scatter plot. 361

List of Appendix Tables

Appendix Table 1. Example loading scores for criteria used in LDA of experimental handaxe assemblage. 359
Appendix Table 2. A sample of LDA scores for handaxes from my experimental assemblage. 360
Appendix Table 3. Sorting table illustrating the classification of a sample of experimental handaxe specimens in comparison to their original, given grouping. 362
Appendix Table 4. Table showing predictive accuracy of LDA by group. 363
Appendix Table 5. Table showing predictive accuracy of 'jackknifed' LDA by group. 364
Appendix Table 6. Sorting table illustrating the classification and 'jackknifed' classification of a sample of experimental handaxe specimens in comparison to their original, given grouping. 365

1. Chapter 1 – Introduction

1.1 Overview

This thesis represents an attempt to produce a methodology which can reasonably and reliably be used to assess the skill-level of knappers in the archaeological record based on the handaxes they left behind. My investigation focuses solely on the quantitative and qualitative analysis of bifaces themselves, and not in relation to the by-products of their creation (i.e., analysis of associated waste material) because a handaxe constitutes the physical outcome of a knapper's ability to attend to multiple complex processes of production. In addition, despite flake debris from production forming a valuable source of data for analysing skill-level, it is very rare that it can be associated confidently with manufactured handaxes in the Acheulean record. Thus, concentrating exclusively on bifaces means that my method can be applied more broadly across Acheulean assemblages than would otherwise be the case.

It must be noted from the outset, that I am not directing my analysis in search of finding and assessing the products of specific, individual knappers in the past. Due to the many factors that act upon the shaping process, recognising the products of specific individuals is very challenging. Indeed, as I show later in my experiments, expert instruction and other influences undermine our ability to accurately isolate and identify the products of separate practitioners. Instead, I shall be appraising archaeological bifaces based on the level of skill they demonstrate along a spectrum of Novice, Intermediate, and Expert. To

discriminate between Palaeolithic bifaces, it was necessary to assemble quantitative and qualitative data from modern experimental knappers of varying abilities who could be regarded as belonging to one of the above skill-level categories. Assessing the aptitude of contemporary knappers was based on the recognised extent of their experience.

As far as I am aware, assessing the skill-levels exhibited in archaeological bifaces has not been previously attempted in detail based on experimental data. This is notwithstanding numerous experimental studies which have been centred on pinpointing the key skill-related behavioural criteria that are observable in handaxe making among existing knappers. The insights from such examinations are often used anecdotally to assign skill-level grades to bifaces and the wider assemblage to which they appertain. Therefore, a methodology which can dependably generate skill-level data from bifaces (excluding associated flake waste) provides a new avenue for collecting data concerning Acheulean knappers and knapping.

I contend that this approach is most useful when utilised at the intra-assemblage level. When applied to a biface assemblage that displays a shared and consistently applied mode of handaxe manufacture, my system has numerous prospective uses. These include, but are not limited to, interpretations of site-function, as well as estimations of the possible broader social makeup of those who contributed to an assemblage. Therefore, I assert that a comparison at the inter-site level between two independent handaxe assemblages that are separated considerably in space and time, and which do not share a coherent method of biface manufacture would be very limited in its scope. Even if the experimental dataset and approach presented in this thesis is used individually at either site to differentiate between bifaces, the validity of evaluations of proficiency between two such groupings

would be reduced when the objectives for production at both sites can be shown to be different. Declaring one biface to be more skilfully made than the other in this scenario is almost meaningless because skill is relative to the pursuit being undertaken and cannot be investigated in a broadscale manner.

Therefore, more useful information can be yielded from my experimental approach when investigations into skill are treated as relative and contextual to the assemblage being examined. Consequently, the results of my experimental analyses are used to assess the ranges of skill-level categories among the consistently shaped bifaces from Gouzeaucourt Level G (GZnivG), the main archaeological focus of this thesis.

1.2 Introducing Gouzeaucourt

GZnivG is one of several archaeological assemblages at Gouzeaucourt, located in the Vallée du Muid in the upper basin of the river Scheldt, northern France (see Figure 1). It has been indirectly dated to MIS (Marine Isotope Stage) 8 (300 – 243 kyr BP), which is considered a complex cold stage within the Early Middle Palaeolithic of north-western Europe. Typically, MIS 8 is characterised by reduced hominin occupation in more northerly regions due to the very cold conditions at this time (Hérisson et al., 2016). Yet, over the course of this period in north-western Europe, considerable technological change occurred (Malinsky-Buller, 2016). At the start of MIS 8, bifaces continued to predominate but by the onset of MIS 7, Levallois and prepared-core reduction seem to have overtaken this (Malinsky-Buller, 2016). Moreover, despite a general trend of increased assemblage variability throughout this stage, Gouzeaucourt biface assemblages interestingly show

atypical morphological and techno-functional consistency over several archaeological levels at the site (Field, 2002; McPherron, 1994; Soriano, 2000).

Level G yielded predominantly ovate bifaces that exhibited a high propensity for tranchet sharpening. Furthermore, reduction in dimension of the handaxes via resharpening was visible and culminated in a wide range of biface sizes. Despite resharpening, the very high technical and typological coherence of GZnivG is repeated in levels H and R and signifies cultural unity between the makers of handaxes at Gouzeaucourt (Soriano, 2000). Further evidence for cultural accord pertaining to biface manufacture in the landscape around Gouzeaucourt is provided by the handaxes from the nearby site of Longavesnes. They closely resemble those from various levels at Gouzeaucourt and are attributed to a similar age (Ameloot-Van Der Heijden, 1993; Tuffreau et al., 2008; McPherron, 1994). However, unlike levels H and R where bifaces were made on flint derived from chalk that lay in a secondary position in nearby slopes (Soriano, 2000), GZnivG bifaces were produced offsite and transported in (Scott & Ashton, 2010; McPherron 1994). This is evidenced predominantly by the absence of any refitting between the bifaces and the many flake tools and lithic debitage found in association with them (McPherron, 1994; Scott and Ashton, 2010; Tuffreau et al., 2008). Soriano (2000) also noted that the only flint nodules worked *in situ* in Level G were of unsuitable morphology and quality for the shaping of the handaxes discovered there. In addition, Soriano (2000) challenged the uniqueness of the deposit at GZnivG due to the identification of some handaxes exhibiting a double patina. Therefore, the assembly of handaxes at Level G most probably accrued from several occupations separated in time (Soriano, 2000). This interpretation renders the regularity of manufacture method seen at GZnivG even more remarkable and is

informative regarding the possible use of the site and the skill-levels of the knappers who contributed to it.

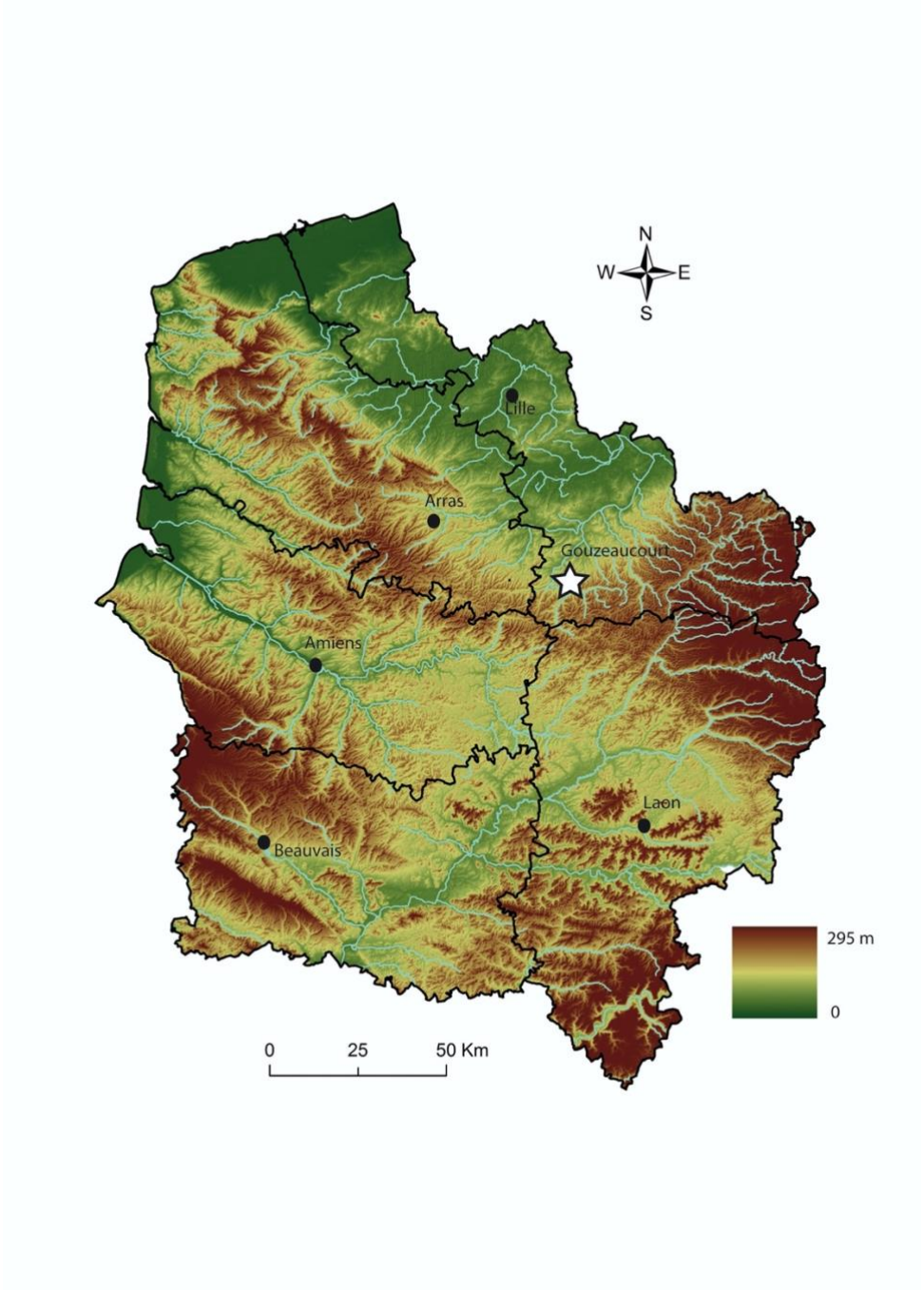


Figure 1. 1. Location of Gouzeaucourt site within the Hauts-de-France region (Nord-Pas-de-Calais and Picardy). Sourced from: Carthage database, courtesy of Dr A. Lamotte, Université de Lille. Created by: P.-G. Salvador, Université de Lille.

Repeated occupation of the site, coupled with the fact that most bifaces were not exhausted in terms of viable resharpening opportunities, suggests that GZnivG was a notable place in the landscape for hominins. The recurrent deposition of allochthonous bifaces suggests Level G acted as a known cache of raw material, the location of which may have been communicated among a group or multiple related groups as they negotiated the increasingly harsh conditions of MIS 8 (Hérisson et al., 2016; Malinsky-Buller, 2016; Scott and Ashton, 2010). In addition, although it is unknown over what distance the handaxes were brought to Level G, the curation of raw material resulting in occasionally small, resharpened bifaces alludes to the possibility that bifaces travelled considerable distances before being left there. With this in mind, I would posit that inexperienced, novice knappers would have likely made only a small contribution to the assemblage. Accordingly, one way of testing this hypothesis is to examine the range of skill-levels displayed by the bifaces within the assemblage. This is something I will do in the ensuing sections of this thesis.

1.3 Approach

My analysis of GZnivG, as already stated, necessitates *inter alia* using the data collected from variably skilled modern experimental knappers. In Chapter 5, I provide a detailed rationale for the materials used in my investigation, and in Chapter 6 I detail the full extent of the analyses involved in my mixed methodology. However, it is necessary to briefly introduce here the key features of my approach before outlining the structure of my thesis in the next section.

The skill-level classification of contemporary knappers is linked to the amount of experience they acquired either before or during the experiments. Using quantitative data from persons with known experience, I have identified thresholds of performance for skill categories (Novice, Intermediate, and Expert) that can be transposed onto and compared with the archaeological data captured from GZnivG bifaces using Linear Discriminant Analysis (LDA). Furthermore, assessments regarding the aptitude of Gouzeaucourt knappers can be strengthened through observing the qualitative commonalities and differences between experimental skill-level categories. This includes, as alluded to above, the effects of factors that influence the levels certain skill categories can reach, both quantitatively and qualitatively (e.g., direct instruction from an expert). As I will show, this data is essential in disentangling overlaps that occur between modern knappers of varying abilities and which can result in misclassifications of bifaces to the wrong skill categories. Moreover, such information is helpful in diluting any ambiguity that may arise when making skill-level assessments of archaeological bifaces. Where ambiguity remains in classifications of skill-level at GZnivG, qualitative analysis was able to provide definitive context and supply resolutions to issues.

Supplementary analyses to the main quantitative and qualitative examinations of archaeological data included the diacritical analysis of GZnivG handaxes. Using observations of flake removal patterns, I generated a further source of quantitative data that helped support determinations of skill-level. In addition, after consultation with my Transfer and Confirmation of Status assessors, I developed a questionnaire for modern flintknappers. It generated ethnographic data (see Catalogue 2) regarding their views on skill related to handaxe manufacture and to knapping more generally. Insights from this are woven into the thesis and used to link methods of analysis and their associated results

to the broader theoretical framework of investigating skill outlined in Chapter 4. This helped me contribute to current discussions, summarised in Chapter 4, related to how we think about skill more generally.

1.4 Thesis Structure

In Chapter 2, I provide a spatiotemporal framework of the Acheulean, outlining which hominins were involved, what behaviours can be attributed to them, and when and where they occurred. I introduce the principal material of the thesis, the biface, and discuss what these artefacts are, and what their evolutionary significance is.

In Chapter 3, I narrow the focus towards providing an overview of how morphological variation of bifaces is a well-trodden discussion that remains contentious. I discuss the various approaches to morphological variation and what have been regarded as the main factors affecting it. Essentially, this boils down to whether variation reflects cultural influences on manufacture or more utilitarian considerations. Here, I agree with Hosfield et al. (2018) and Wynn and Gowlett (2018), who challenge a conservative and variably same (e.g., Isaac, 1977) reading of the Acheulean because they consider this interpretation to homogenise the variation within it. Furthermore, I contest previous approaches that deem handaxe shape to be unrelated to the morphological intentions of the knapper. I suggest that morphology is deliberate and go on to provide the background for introducing the notion of skill as a relevant topic in the debate (i.e., that finding common goals are useful for investigating skill).

In Chapter 4, I deliver a discussion of the theoretical approaches to skill that have influenced my thinking. The purpose of this chapter is to supply a framework with which to interpret the results presented in Chapter 7.

I provide a rationale for experimental and archaeological materials used in my thesis in Chapter 5, prior to outlining how they will be analysed in the following Chapter (6). Most of this chapter is focused on stipulating sufficient evidence of the suitability of GZnivG for investigating skill-level differences. This comprises a full background of the assemblage and the wider site, as well as contextualising it within the wider contemporary Acheulean landscape. Moreover, I supply further evidence to support work previously carried out by McPherron (1994), Soriano (2000), and Field (2002) regarding the morphological and techno-functional consistency observable at Gouzeaucourt.

As previously mentioned, Chapter 6 is my Methodology chapter. I outline here each analysis undertaken, as well as contribute further rationale for why they are being performed. Furthermore, I present my hypotheses for both experimental and archaeological assemblages in relation to the methods of analysis described.

This is followed, in Chapter 7, by a full presentation of my quantitative and qualitative results for both experimental and archaeological investigations. This chapter concludes by summarising my findings and relating them back to the hypotheses.

Chapter 8 is the concluding chapter, where key findings are summarised and discussed. It also considers where the methodology could be improved by re-testing it against a sample from another Palaeolithic site at Boxgrove. Moreover, I examine the wider

contribution of the results to the study of skill using the theoretical discussion outlined in Chapter 4. Finally, new directions for productive future study are stipulated.

In summary, my thesis aims to use data yielded from a mixed methods analysis of experimentally made bifaces to discriminate between bifaces found in Gouzeaucourt Level G by their perceived skill-level. While I am using GZnivG as an initial case study to test my method of approach, the wider aim of this thesis is to present a more generalised experimental method for yielding new data from the Acheulean record, one that reasonably and reliably discriminates between bifaces via skill-level and can be used on any biface-rich assemblage that exhibits consistency in approach to manufacture – and where the associate waste material is absent. In addition, I aim to contribute to broader theoretical discussions of skill in the Acheulean, and potentially beyond this period.

2. Chapter 2 – The Acheulean

2.1 First principles: who, what, where, when?

The term '*The Acheulean*' refers to a technology, also known as Mode II *sensu* Clark (1969), represented by bifacially worked stone tools manufactured by hominins in the Lower and Middle Pleistocene. It was first defined by Gabriel de Mortillet (1873) who identified bifaces as the principal stone-tool type of the Acheulean technocomplex (de La Torre, 2016). de Mortillet (1873) attempted to define Palaeolithic cultures based on characteristic stone-tool types, naming periods after the first, or most typical, site where tools were discovered (de La Torre, 2016). The eponymous site was located in St Acheul (hence, Acheulean), in the Somme Valley, northwestern France (de La Torre, 2016; de Mortillet, 1873).

Since de Mortillet's time there has been considerable debate regarding the geochronological framework of the Acheulean and its relation to the earlier core and flake industries (de La Torre, 2016; de Mortillet, 1873). A comprehensive detailing of this debate is beyond the remit of this introductory background, however, de Mortillet's (1873) view of the biface as the defining stone-tool type of the Acheulean still exists today (de La Torre, 2016). Acheulean bifaces superseded simpler core and flake technologies, grouped as Oldowan or Mode I technology, but did not fully replace them. Instead, they added another dimension to the hominin toolkit (de La Torre, 2016). Currently, there is now a notional split between 'earlier' and 'later' Acheulean bifaces, those dated to before 600 kyr BP deemed generally less symmetrical, thicker, with fewer removals and more sinuous cutting edges, whereas those later being described as more

'refined' due to greater degrees of congruent symmetry, thinner profiles and straighter edges (see Figures 2.1 and 2.2) (Iovita et al., 2017; Schick and Toth, 2017; Shipton, 2018; Wynn, 2002).

The earliest occurrences of Acheulean bifaces in the archaeological record date to around 1.8 million years ago in eastern Africa (Beyene et al., 2013; Lepre et al., 2011; Shipton, 2018). The Acheulean technocomplex covered a vast spatio-temporal range, persisting for over 1.5 million years (henceforth shortened to 'myr', or 'kyr' for thousands of years) throughout Africa, the Middle East, Asia and Europe (de La Torre, 2016; Shipton, 2019, 2018). The biface acts as the index fossil for this technology, with *Homo erectus sensu lato* and *Homo heidelbergensis* most commonly considered responsible for its manufacture. *Homo erectus sensu lato* was the first to make and use bifaces in Africa around 1.8 myr BP and *Homo heidelbergensis* is attributed to making them from around 700 kyr BP, refining and advancing bifacial technology further (Iovita et al., 2017; Schick and Toth, 2017; Shipton, 2019, 2018; Stout et al., 2014; Wynn, 2002).

2.2 What is a biface?

The term 'biface' is most widely used to refer to all Acheulean forms of large cutting tools (LCTs), including picks, cleavers (see Figure 2.3) and handaxes (de la Torre, 2016). However, de la Torre (2016) argues that in the early African Acheulean LCTs can often be unifacial, trihedral or quadrangular. Therefore, he prefers to use the term 'handaxe' when referring to bifacially worked Acheulean stone-tools. In the later Acheulean of Middle Pleistocene Europe, the spatiotemporal focus of this thesis, unifacial LCTs are rare, cleavers still exist as a recognisable type, but handaxes are most common.



Figure 2. 1. Planform and profile views of Early Acheulean handaxe, Kokiselei, Kenya. (Source: Texier, 2018). Photograph: P.-J. Texier MPK/WTAP.

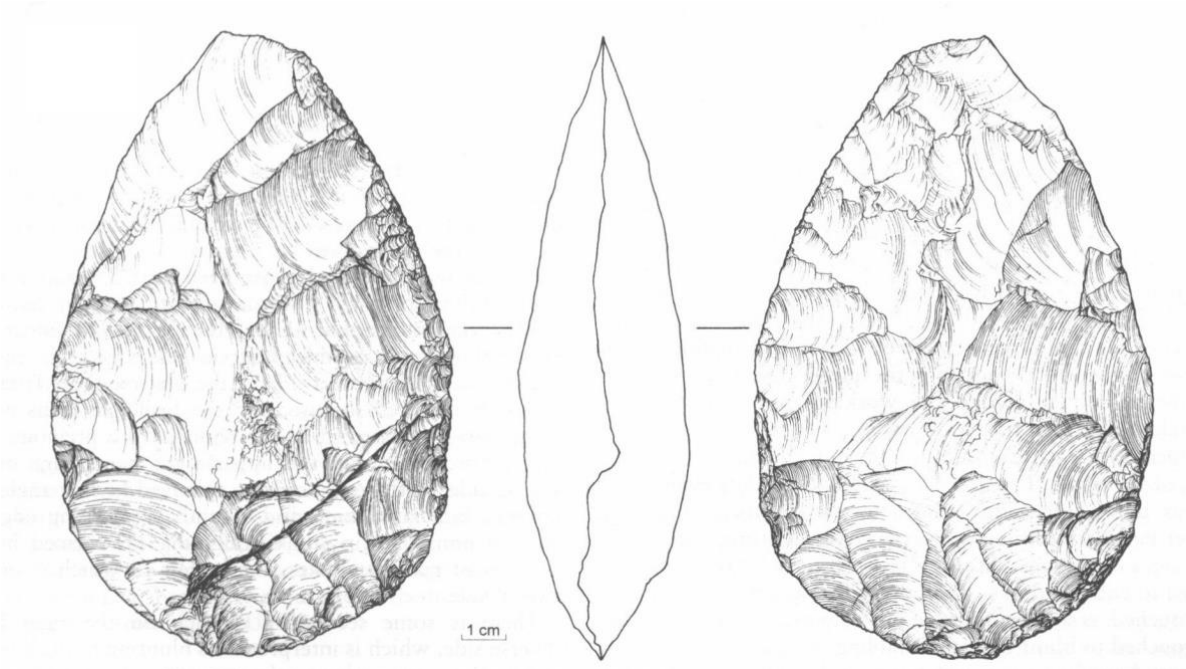


Figure 2. 2. Planform and Profile views of a Later Acheulean handaxe, Boxgrove (no. 150 from Unit 4C), UK. Central line in profile view denotes edge. (Source: Roberts et al., 1997).

Arguably, whilst the term ‘biface’ still encompasses handaxes and cleavers in later Acheulean contexts, it continues to contribute to the maintenance of a definition which is too broad. It subsumes too much variation among stone-tool types over vast spatio-temporal contexts (where the principal link is that they have two faces). However, categorising types of bifaces into handaxe or cleavers et cetera, still does not adequately cover the variation that exists within each of these categories. Therefore, there will always be a need to contextualise and describe the type of biface, or handaxe, one is discussing. This is because either term does not code for a particular unchanging template but variations on a common, basic template.

For example, both English and French traditions of Acheulean research have specific nomenclature for the shapes and general forms of particular handaxe types (Bordes, 1968; Roe, 1968; Wymer, 1968). Therefore, whether using ‘biface’ or ‘handaxe’, one still needs to qualify the typology of the tool. It also suggests that arguing over whether ‘biface’ or ‘handaxe’ as terms are too broad, or not broad enough, is inconsequential to the ultimate understanding of an artefact, as researchers will continue to make the necessary comparisons and debate about the place of the artefact as an actual category, as well as its place within the Acheulean.

To illustrate this point, White’s (2006) query of the veracity of the term ‘cleaver’ (Figure 2.3) may shed some light. The principal differentiating feature between handaxes and cleavers relates to the shape of their tips and, as a result, their perceived function. Cleavers exhibit broader and squarer tips than handaxes which exhibit a wider range of proximal apices along the spectrum of rounded to pointed (see Figure 2.4) (Roe, 1968; White, 2006). The postulated difference in use between the two forms is related to tip shape, with

the broader, squarer tip of the cleaver being perceived as more conducive to chopping whilst handaxe tips are better designed for slicing. However, White's (2006) analysis of the British Acheulean record proposes that the rarity of cleavers may undermine their status as a *bona fide* bifacial tool category. In fact, he suggests they may not represent the continuation of a stone-tool type from earlier in the Acheulean but may represent the unsuccessful performance of a novel Acheulean handaxe technique – the 'tranchet' flake (see Figure 2.5). Therefore, definitions of tool types within Acheulean categories are subject to change. As such, the position of the cleaver within different Acheulean contexts is undecided – is it a category of Acheulean biface or is it to be subsumed into the profusion of varieties of Acheulean handaxes?

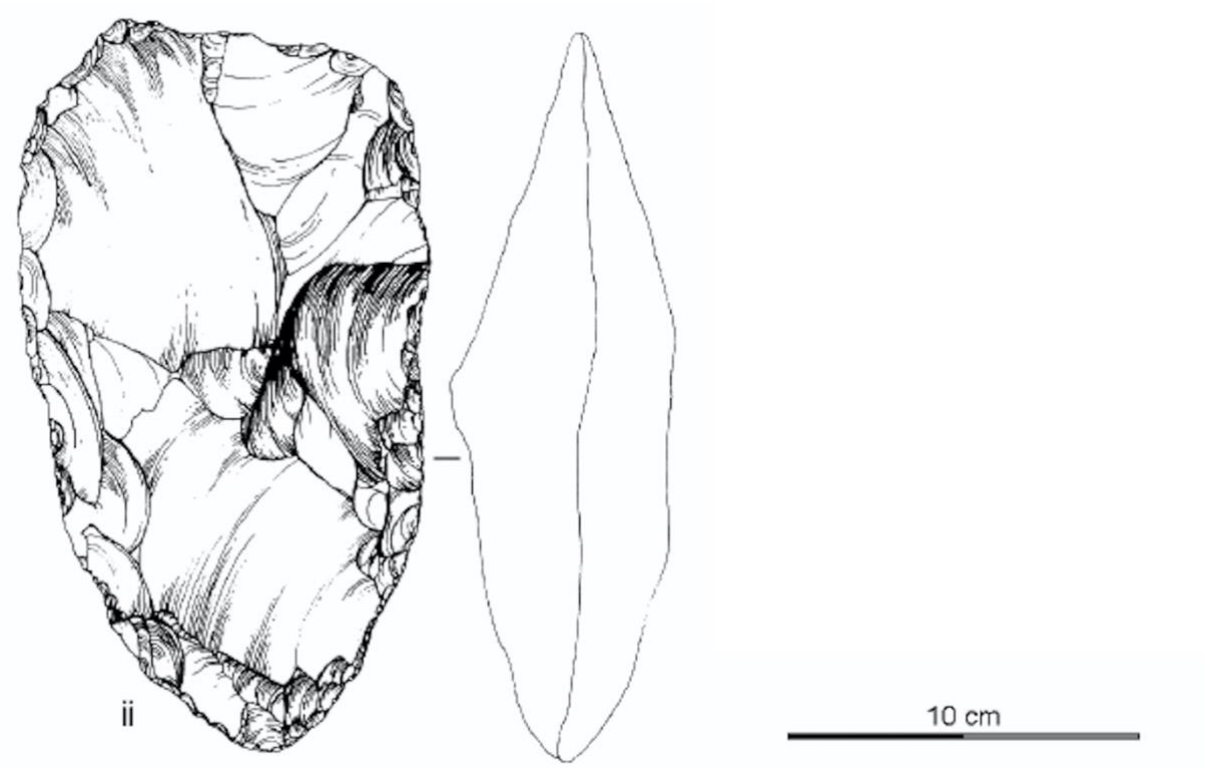


Figure 2. 3. Planform and Profile view of a Cleaver from Highfield, UK. (Source: Wenban-Smith, 2001).

What can be agreed upon, however, is that Acheulean bifaces, or handaxes, share common features that can link them under either heading. Handaxes exhibit two distinct faces separated by a plane of intersection in profile view which exhibit deliberate flaking of variable intensity on both. This intersection is the edge at the point where the two faces meet extending around most or all the tool. Additionally, there is a distinction between the butt and tip – the tip is sharper and narrower and where the cutting function of the biface/handaxe was focused and the butt is blunter and easier to handle. Lastly, bifaces also demonstrate variable symmetry in both plan and profile views (Shipton, 2018).

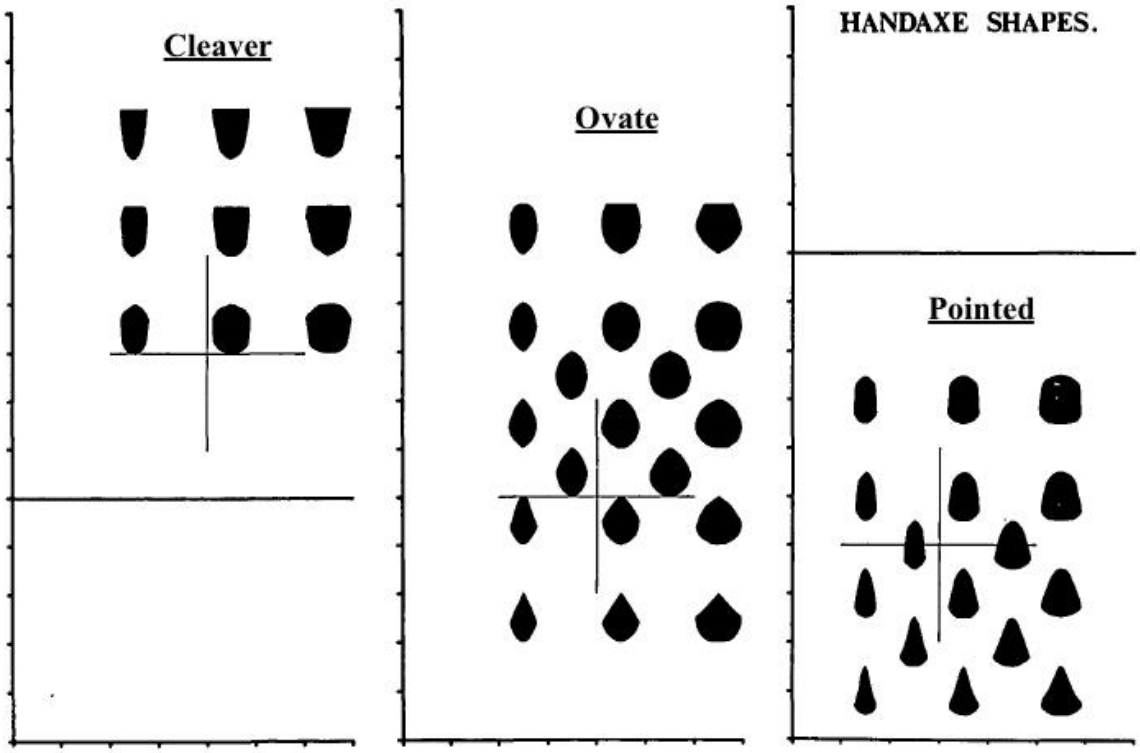


Figure 2. 4. Range of handaxe shapes (tip shapes) as defined by Roe (1968). (Source: Roe 1968).

Gowlett (2006) provides more detail with his list of 'design imperatives' (GLOBFELTS) which, he suggests, guided knappers and emphasized the way handaxes, or bifaces, were

intended to work (Wynn and Gowlett, 2018). GLOBFELTS consists of six design imperatives shared relatively consistently by handaxes, or bifaces, in the Acheulean:

- Glob-butts (the centre of gravity is situated toward base for enhanced ergonomics);
- Forward extension (longer tools could provide purchase via an extended cutting edge but with the centre of gravity remaining in the hand, hence thinner tips);
- Support for the working edge (via bifacial trimming from the margin constituting the plane of intersection between both faces onto either face, subsequently producing the lenticular cross-section common to Acheulean handaxes);
- Lateral extension (another ergonomic consideration used to compensate for longer tools being more unbalanced);
- Thickness adjustment (used to control weight and edge angles for potentially different uses - thicker handaxes maybe having different functions to thinner ones);
- Skewness (handaxes may have been biased towards handedness).

  (Wynn and Gowlett, 2018:22).

Although not all handaxes, or bifaces, exhibit all of Gowlett's (2006) design imperatives in each instance, what GLOBFELTS provides is a baseline for understanding Acheulean bifaces. Some may not exhibit true bifacial working, particularly developed symmetry or effectively thinned cutting edges, but there are commonalities among Acheulean handaxes, outlined in the above definitions, that preclude the inclusion of other Acheulean tools such as picks or polyhedral LCTs (de la Torre, 2016).

Therefore, this definition highlights the necessity of a common term for tools that exhibit striking commonalities over time and space. Conversely, a collective term is simultaneously insufficient for describing the variation that exists within it. Proper description and analysis of individual biface types, subsumed within the term, is required. Indeed, each of the characteristics outlined above, can, and have, been achieved in different ways over the course of the Acheulean. This manifests most often in visible traces of morphological variation or variability in method of manufacture.

Therefore, I will use the terms 'handaxe' or 'biface' interchangeably at a broader level of discussion but will caveat each term with reference to their particularities when discussing morphological or other differences – e.g. ovate versus pointed, or whether made via debitage or façonnage (Coolidge and Wynn, 2009; Roe, 1968; Sharon, 2010; Shipton, 2018; White, 1998). It is these differences in morphology and method of manufacture, as well as variations on the characteristic handaxe definition, that are highly significant for our understanding of their use, the social constraints governing their manufacture, and the complexity of production. Combined, they lead us to a greater comprehension of the evolutionary and social significance of the Acheulean handaxe.

2.3 Evolutionary Significance of bifacial technology

Acheulean handaxe technology signals an important landmark in human evolution, heralding a development in the cognitive and behavioural complexity amongst the hominins responsible for its invention and subsequent innovation. For the first time, pre-modern humans were shaping their rocks towards an intended aesthetic end form, a

behaviour requiring increased planning abilities and cognitive power, in general (Muller et al., 2017; Stout et al., 2015). Moreover, hominin cognitive advancements occurred within more socially complex and highly communicative groups who shared intentions, learned from each other and were influenced by social conventions which they helped to replicate (Shipton, 2010; Shipton and White, 2020).

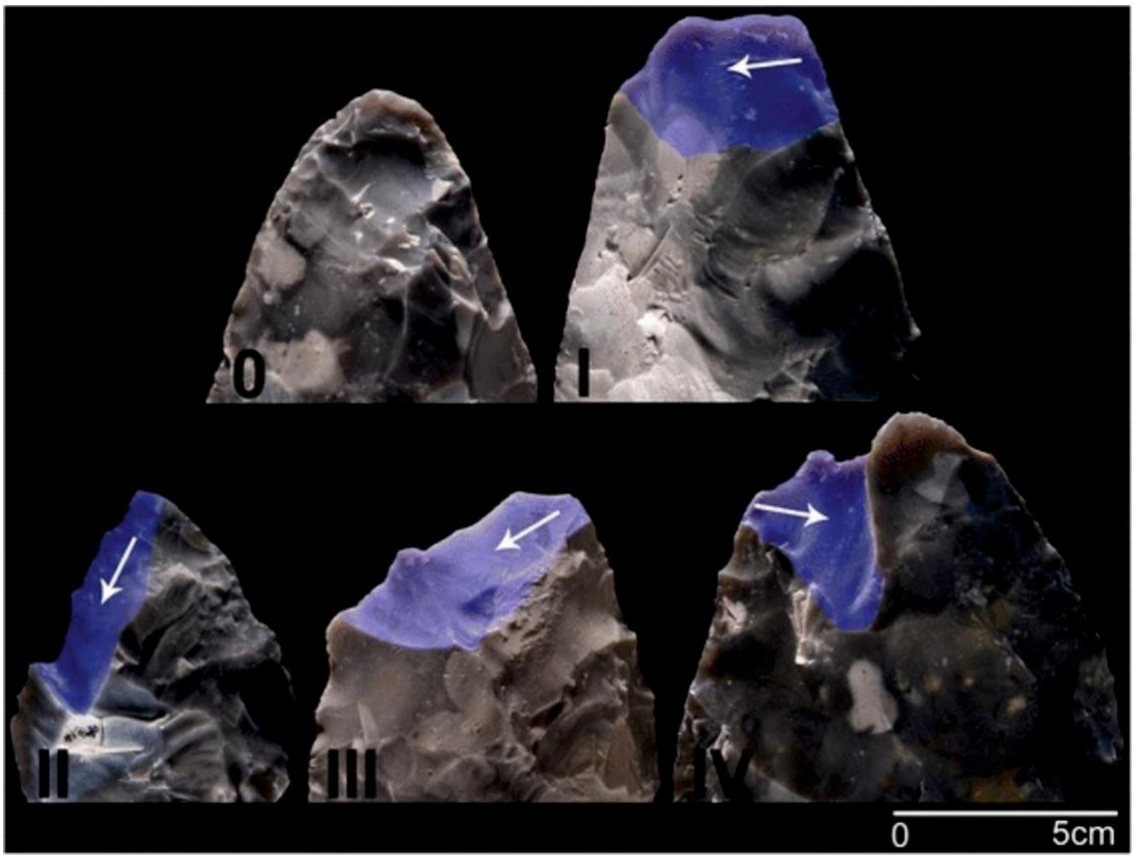


Figure 2. 5. Distal (tip) ends encountered at Boxgrove: type 0, no tranchet removals; type I, distal tranchet: a transverse blow producing a straight or convex tip; type II, lateral-distal tranchet: longitudinal blow long lateral edge and shorter distal (tip) edge; type III, invasive tranchet: often deriving from successive phases of oblique tranchet retouch, or perhaps an overshooting of intended outcome; type IV, unsuccessful tranchet: produced when force does not go far enough across the face of the piece and ends with breakage of distal tip. (Source: García-Medrano et al., 2019).

2.3.1 Cognition

Experimental studies using functional brain imaging techniques on modern knappers have helped highlight the potential cognitive developments associated with the Acheulean (Stout *et al.*, 2015). Stout *et al.* (2015) found that areas of the brain central to conceptual knowledge, such as prediction and simulation of accurate strategic judgements, show increased activation and connectivity for Acheulean technological stimuli than for initial Mode I core and flake (Oldowan) technological stimuli. This suggests handaxe technology required brains capable of dealing with greater predictive and strategic complexity, as well as superior abilities in visualising discrete stages of manufacture (Högberg et al., 2015; Muller et al., 2017; Stout et al., 2015).

A primary innovation of Acheulean bifacial manufacture was the ability to remove large flakes to be used as tool blanks (Sharon, 2010). This debitage method developed in complexity towards the later Middle Pleistocene with the innovation of prepared core technologies, Victoria West (Li et al., 2017) and Kombewa methods (Herzlinger et al., 2017). A second innovation, more common outside of Africa, was the façonnage reduction of an entire stone nodule to produce the handaxe (Coolidge and Wynn, 2009).

A key later Acheulean innovation was the integration of platform preparation and subsequent removal of invasive flakes, normally using an organic soft hammer (Roberts and Parfitt, 1999; Stout et al., 2014). This innovation enabled the production of much thinner and more refined (thinner relative to breadth and generally more symmetrical) tools as well tranchet tip removal (Fig. 2.5). Together these innovations are argued to

evidence the intentional imposition of a mental template because they demonstrate the ability to perform an adaptive set of hierarchically organised behaviours based on complex causal understanding (Garcia-Medrano et al., 2019; Gowlett, 1984; Pelegrin, 2005, 1993; Wynn, 2002, 1995). Additionally, the human fossil record concomitant with the innovations in Acheulean stone-tool technology points, although indirectly, towards a significant cognitive and behavioural developments among handaxe producing hominins.

A dramatic increase in brain and body size, as well as a reduction in post-canine dentition, is observable in *Homo erectus* and then *Homo heidelbergensis*. Among *Homo erectus* there is selection towards an increase in brain size over the course of the Acheulean, with early specimen cranial capacity measuring around 600-800cm^3 and increasing to well beyond 1000cm^3 in later specimens (Van Arsdale, 2013). *Homo heidelbergensis*, at 1,250cm^3, pushes this further with a brain size nearer that of Neanderthals and *Homo sapiens* (Bae, 2013; Rightmire, 2008, 2004). Enlarged body size among *Homo erectus* onwards is evidenced by the adolescent Turkana boy who was already over 5 feet (152.4 cm) tall at around 1.5-1.6 myr BP (Van Arsdale, 2013; Walker and Leakey, 1993). Moreover, relative to australopiths, *Homo erectus* onwards possessed smaller post-canine dentition (premolars and molars). The molarization of their premolars are much reduced and the mandible is much more gracile (Van Arsdale, 2013).

These skeletal, cranial, and dental transformations would appear to coincide with the beginnings of Acheulean technology and its association with butchered faunal remains. It suggests a behavioural shift towards greater consumption of meat and a movement away from a highly fibrous plant-based diet. This is supported by toothwear analysis

(Ungar et al., 2006) and stable isotope examination of *Homo erectus* teeth (Ungar and Sponheimer, 2011). However, debate remains regarding when exactly hunting became a staple of the human behavioural repertoire in addition to scavenging (Van Arsdale, 2013). A comprehensive overview of this debate in the earlier Acheulean is not pertinent to this thesis, but by the later Acheulean, the period I will be focusing on, hunting was certainly part of the hominin behavioural repertoire.

Later Acheulean sites such as Gesher Benot Ya'aqov, Israel, dated to 700-750 kyr BP (MIS 18) (Alperson-Afil et al., 2009; Goren-Inbar et al., 1994; Rabinovich and Biton, 2011) and Boxgrove, Britain, circa 500 kyr BP (MIS 13) (Pope, 2002; Pope et al., 2020a; Roberts and Parfitt, 1999), unequivocally demonstrate the hunting of medium to large fauna. This helps maintain the view that not only can bifacial technology be used as a marker for increased behavioural complexity in *Homo erectus* but that advances in this technology can potentially be indicative of wider behavioural developments. Indeed, progressions such as hunting are added to in the later Acheulean with respect to fire use, at Gesher Benot Ya'aqov (Alperson-Afil et al., 2018, 2017; Goren-Inbar et al., 2004), and clear evidence for the use of hunting weaponry at Schöningen (Thieme, 1997) and Boxgrove (Milks et al., 2019; Roberts and Parfitt, 1999).

Innovations in controlled use of fire also occur during the Acheulean. Wrangham (2010) suggests that cooking dates as far back as ~1.7 myr BP, and Gowlett et al. (2012) note that fire was a major driver in the dietary changes required to permit sufficient encephalization of such an energetically expensive organ as the brain. This, in turn, led to further technological and social benefits – i.e., more efficient tools and weaponry. Debate around the validity of evidence for controlled fire in the earlier Acheulean at

Chesowanja and Koobi Fora ~1.4-1.6 myr BP and Swartkrans and Wonderwerk, South Africa, at ~1.0 myr BP remains. However, from the later Acheulean onwards, evidence for controlled use of fire (Gesher Benot Ya'aqov and beyond ~0.7 myr BP) is more widely accepted. This more advanced behavioural repertoire was more intensely social and underlines how cognitive advancements of Acheulean hominins permitted greater and more complex sociality (Berna et al., 2012; Goren-Inbar et al., 2004; Gowlett et al., 2012).

Around 600 kyr BP there was a notable increase in brain size in *Homo heidelbergensis*, as well as marked innovations in handaxe manufacture and more concrete examples of controlled fire use extending into Eurasia (Gowlett et al., 2012). Therefore, analysis of the Acheulean lithic record, in conjunction with associated fossil, faunal and other behavioural records can demonstrate developments in hominin cognition over the course of the Acheulean via subsequent innovations.

2.3.2 Social Learning

The complexity of handaxe manufacture is significant for assessing hominin social systems of learning, especially given the very wide spatio-temporal extent of the Acheulean. Despite recent contentious calls to view handaxe manufacture as a behaviour partly under genetic control (Corbey, 2020; Corbey et al., 2016), it is clear, as Hosfield et al. (2016) and Wynn and Gowlett (2018) contend, that the Acheulean technocomplex is more reliably explained by sociocultural transmission.

Experimental observations of biface manufacture, and ethnoarchaeological studies witnessing stone-tool manufacture, illustrate that, despite calls to the contrary (Corbey *et al*., 2016), biface manufacture is not an instinctive behaviour genetically determined and easily grasped (Geribàs et al., 2010; Pargeter et al., 2020, 2019; Shipton, 2018). One needs to learn how to knap a handaxe effectively through repeated and prolonged practice towards a known goal (Shipton and Nielsen, 2018). Declarative, or theoretical, knowledge of the procedure is insufficient for knowing what the process entails, never mind realising that process in stone (Pelegrin, 1990; Shipton and Nielsen, 2018; Stout et al., 2015). Practical know-how, or embodied knowledge, is developed through repeated engagement with the raw material towards an objective (Malafouris, 2010; Pelegrin, 1990). Learning from mistakes is crucial. As Pargeter *et al.* (2020, 2019) emphasise, it takes grit and determination on the part of the knapper to learn and master handaxe manufacture. Such an illustration has implications for how we envisage and interpret hominin learning capabilities and processes in relation to biface manufacture.

At the very least, replication of a basic bifacial handaxe template, one which persisted with variation for over 1.5 myr (Shipton, 2018, 2019), presupposes hominins possessed theory of mind and could decode one another's intentions (Gärdenfors, 2017; Kline, 2015). However, this does not explain how they understood the complexity of the operating chain of manufacture that ultimately produces a handaxe. Lycett (2019) maintains that imitation alone could produce the patterns we see in the Acheulean. However, the difficulty of practice and length of skill acquisition via observation and imitation alone is inefficient and lengthy (Muller et al., 2017). In fact, Shipton (2019) suggests that imitation, and even over-imitation, would have been insufficient to transmit difficult behaviours, such as platform preparation, seen in Later Acheulean handaxes.

This introduces the possibility of more complex transmission methods, such as direct instruction and teaching, being employed by Acheulean knappers, perhaps in addition to imitation (Lycett, 2019). Indeed, Morgan et al.'s (2015) experimental investigation of the effectiveness of different social transmission mechanisms (reverse engineering, imitation/emulation, basic teaching/facilitated observation, gestural teaching, and verbal teaching) demonstrated that, across multiple quantifiable measures, transmission, and consequently the ability of the learner to perform more skillfully, improves with purposeful teaching. This is especially apparent with language, but not with imitation or emulation (Morgan et al., 2015). With their findings they suggest that teaching or proto language may have been necessary for the innovation of Acheulean biface technology. However, as Lycett (2019) points out, a range of social transmission methods may occur simultaneously in one learning context.

It is not within the remit of this thesis to present here a comprehensive synthesis of debates pertaining to whether Acheulean handaxe-making hominins possessed language abilities that aided knowledge transference and skill acquisition in learners. However, Stout (2008) and Stout and Chaminade (2012, 2009) highlight that networks for speech and object manipulation overlap in the brain and further explain that both language and production of Mode II tools are hierarchically structured processes. They use this to speculate on evolutionary origins of language, proposing a 'technological pedagogy' hypothesis, which explains how intentional teaching provided the adequate 'scaffold' for the evolution of intentional vocal communication.

Alternatively, Barham and Everett's (2021) Peircean approach defines language as 'communication based on symbols rather than grammar'. They argue that the presence of

symbol-based language among Acheulean handaxe-making groups is evident in the normative social conventions that are expressed when bifaces are made a certain way (when there are multiple ways to make them). Barham and Everett (2021) note that formation of conventions requires understanding and sharing, via social transmission, of abstract cultural knowledge and arbitrary meaning that can be then applied to the creation of a material object (i.e., a handaxe).

Overall, whether Acheulean hominins possessed language (or not) is an important question but one that is still not resolved. Some argue for language's deep origins (Barham and Everett, 2021; Stout and Chaminade, 2012), whereas others view the emergence of language as a more recent and purely *Homo sapiens* ability (Klein, 2009, 2000). Experimental studies suggest that acquisition of handaxe manufacturing skill within a learning environment would be more efficient amid contact with an expert (Shipton and Nielsen, 2018) due to the complexity of Acheulean manufacture (Muller et al., 2017). In addition, Kempe *et al.* (2012) and Schillinger *et al.* (2014) show that handaxe manufacture is particularly prone to copying error but the range of shapes produced in the archaeological record is limited. This is particularly true at the intra-assemblage level, where considerable homogeneity of biface form is visible at Boxgrove and elsewhere. Schillinger et al. (2015) suggests that imitation reduces copying error because it entails copying another's actions and not just their final products (Wynn and Berlant, 2019). Corbey et al. (2016) suggest this homogeneity has a genetic explanation, and is not the result of social transmission. However, as Shipton (2019) suggests, Corbey et al. (2016) disregard the learnt aspect of engaging with the process and becoming skilled in notably difficult tasks such as platform preparation. Therefore, the complexity of the task, combined with the requirement for prolonged practice, and handaxes within

assemblages like Boxgrove displaying shared objectives of manufacture (García-Medrano et al., 2019; Leroyer, 2016; McNabb, 2017; Shipton and White, 2020; Wynn and Berlant, 2019), lead me to agree with Morgan et al. (2015). Basic teaching (facilitated observation) is the minimal plausible form of social transmission for Acheulean biface manufacture. Gestural instruction is probable but some form of verbal instruction is also possible.

What is more certain than accurately demonstrating the method of social transmission used by Acheulean hominins, however, is Whiten's (2015) proposal that knapping skill is propagated via a 'helical curriculum' of social learning. In Whiten's curriculum, skill acquisition occurs in cycles of knowledge transference and subsequent practice within a social scaffold providing the necessary motivation for learners to continue learning (Botha, 2015; Stout, 2002; Terashima, 2013). An ethnographic example of this is provided by Stout's (2002) analysis of the Langda adze-makers of Indonesian Irian Jaya, where he observed extant craftspeople learning the process of adze production within such a scaffold. Novices imitated Experts, and Experts intervened in Novice learning to maintain and enhance learner determination when inevitable frustration set in (Stout, 2002). Indeed, Pargeter et al.'s (2019) modern experimental research demonstrates that cultivating a learning environment which maintains learner motivation is critical to the development of Acheulean handaxe manufacturing skill among modern human participants. Although these are modern examples and cannot be considered as direct evidence for what occurred in the Acheulean, evidence of consistent manufacture method and style over time among Acheulean handaxe makers (e.g., at Boxgrove) suggests knowledge transference and skill development occurred within social, structured systems

(García-Medrano et al., 2019; Hutchence and Scott, 2021; Leroyer, 2016; McNabb, 2017).

Therefore, experimental research confirming that Acheulean handaxe manufacture was sufficiently difficult to require a prolonged learning period, as well as potentially purposeful instruction of beginners by those more experienced, is highly significant. Firstly, support of learners in acquiring handaxe knapping skill may highlight for the first time a direct pedagogical dynamic in our hominin story over mere observation and imitation (Morgan et al., 2015). In addition, considerable learning time means there should be a visible signature that shows a difference between those learning and those already skilled (Bamforth and Finlay, 2008; Finlay, 2015, 2011; Shea, 2006). Demonstrating this difference in skill-level is the basis of this thesis.

2.3.3 Utilitarian vs. non-utilitarian significance

It is widely accepted that Acheulean handaxes served some important practical, utilitarian purpose. Evidence of handaxe manufacture at Early and Middle Pleistocene kill sites and associated use-wear and residue analysis (Keeley, 1977; Mitchell, 1998; Solodenko et al., 2015) suggests they were tools for processing animal carcasses. In addition, others have suggested they could be used for a variety of other functions such as plant cutting or woodworking (Dominguez-Rodrigo et al., 2001) or a combination of purposes (like a Swiss Army knife (Schick and Toth, 1993)). However, the most widely agreed upon use due to their size and weight was as tools for butchering hunted and/or scavenged fauna.

To be effective as butchery tools, handaxes needed to be sharp enough to cut through fat, muscle, sinew, and cartilage as efficiently as possible. Therefore, creating a useful handaxe is not as simple as bifacially shaping a piece of stone, but relates to the success of the tool from a practical perspective. It requires its maker to perform the difficult knapping gestures (inherently related to the defining characteristics of the handaxe outlined above) skillfully and accurately. This is necessary to produce a serviceable handaxe.

The overall morphology of the handaxe and its relative practicality relate to the ability of the knapper to create and manage the plane of intersection between the two faces, thin the piece effectively, establish planform and profile symmetry, as well as create thin, straight, and sharp cutting edges allowing for the efficient butchery of large mammals (Shipton, 2018). Successful completion of the above, results in an evenly weighted, thin, and simple to handle handaxe which can cut more easily (Shipton, 2018). Each of these behaviours requires a greater appreciation of fracture mechanics and increased striking accuracy by the knapper using a percussor, highlighting the increase in behavioural complexity from earlier Oldowan, core and flake technologies as outlined above (see sections 2.2.1 & 2.2.2).

However, studies like Stout's (2002) introduced above, raise questions regarding the potential for non-functional, yet socially significant, roles stone-tools could have played in hominin life. The Langda imbue raw material with intentionality, perceiving them as living subjects that are believed to grow and age as people do (Coward and Gamble, 2010; Stout, 2002). Moreover, care must be taken to avoid angering pieces through improper practice (Coward and Gamble, 2010; Stout, 2002). It is clear then, that for the Langda,

social relations and identity are critical components of stone-adze production. The formal apprenticeships, that take place over several years, are a method for not only reiterating that there is a correct way of doing things but repeating that through maintaining correct procedure one maintains a link to the past and ancestral ways (Coward and Gamble, 2010; Stout, 2002).

The symbolism involved in the manufacturing process of stone-adzes among the Langda must, of course, be qualified in that it exists in a modern semi-sedentary society (Stout, 2002). The Langda do not make Acheulean handaxes and they are a completely different species from those who did. Other caveats exist, including whether Acheulean hominins had the capacity for symbolism (Barham and Everett, 2021; Gärdenfors, 2004)? However, the key point to take from examples like the Langda is the potential for non-functional, socially significative importance to be attached to utilitarian items, which could reasonably include Acheulean handaxes.

Indeed, significative features of handaxes have long been proposed. The majority focus on the high degree of planform symmetry in some handaxes and occasional visibly deliberate attempts to correct planform symmetry, as seen at Elveden and Boxgrove in Britain (Ashton and White, 2003; Pope et al., 2006). Kohn and Mithen's (1999) prominent 'sexy handaxe hypothesis' proposed that symmetry was used to signify reproductive fitness. However, McNabb and Cole (2015) argue symmetry, and any significative meaning attached to it, is likely to be situational and/or local. This view is strengthened by the added observation that handaxes were often made quickly for expedient use around kill sites (McNabb et al., 2018). Indeed, given that an experimental study by Machin *et al.* (2007) was inconclusive as to the impact of planform symmetry

in practical butchery activities, it may be that White & Foulds (2018) and Hodgson (2015, 2009) are correct and, although symmetry is intentional, it may not be significative of anything other than it gave hominins pleasure and subsequently became part of hominin social practice (Wynn & Gowlett, 2018). In support of this notion, planform symmetry was not the sole preserve of skilled individuals – even poorly made handaxes are and can be symmetrical (White and Foulds, 2018; Winton, 2005).

Currently, therefore, we cannot yet glean the specific meanings behind any social significance that may be attached to handaxes (Wynn and Berlant, 2019). However, deliberate imposition of symmetry is a significant step in the human evolutionary story, despite inconclusive practical reasons for its presence.

More convincing evidence for non-utilitarian considerations behind handaxe manufacture is demonstrated by morphological homogeneity of assemblages of associated handaxes in comparison to the wider Acheulean morphological variability. For example, multiple scholars (García-Medrano et al., 2019; Leroyer, 2016; McNabb, 2017; Shipton and White, 2020; Wynn and Berlant, 2019) note the marked similarity of handaxe shape and characteristics produced at Boxgrove, c. 500 kyr BP (see Figures 2.6 and 2.7). This perceived homogeneity of form has led McNabb (2017: 405) and Garcia-Medrano et al. (2019: 416) to argue that Boxgrove knappers were reproducing a ‘relatively standardised template’ that circumvented raw material differences to achieve a ‘common goal’ that was ‘set in the knapper’s mind’. In theory, this shared and consistently applied handaxe design could be significative of group membership (Hutchence & Scott, 2021; McNabb, 2017; Garcia-Medrano et al., 2019; Leroyer, 2016;).

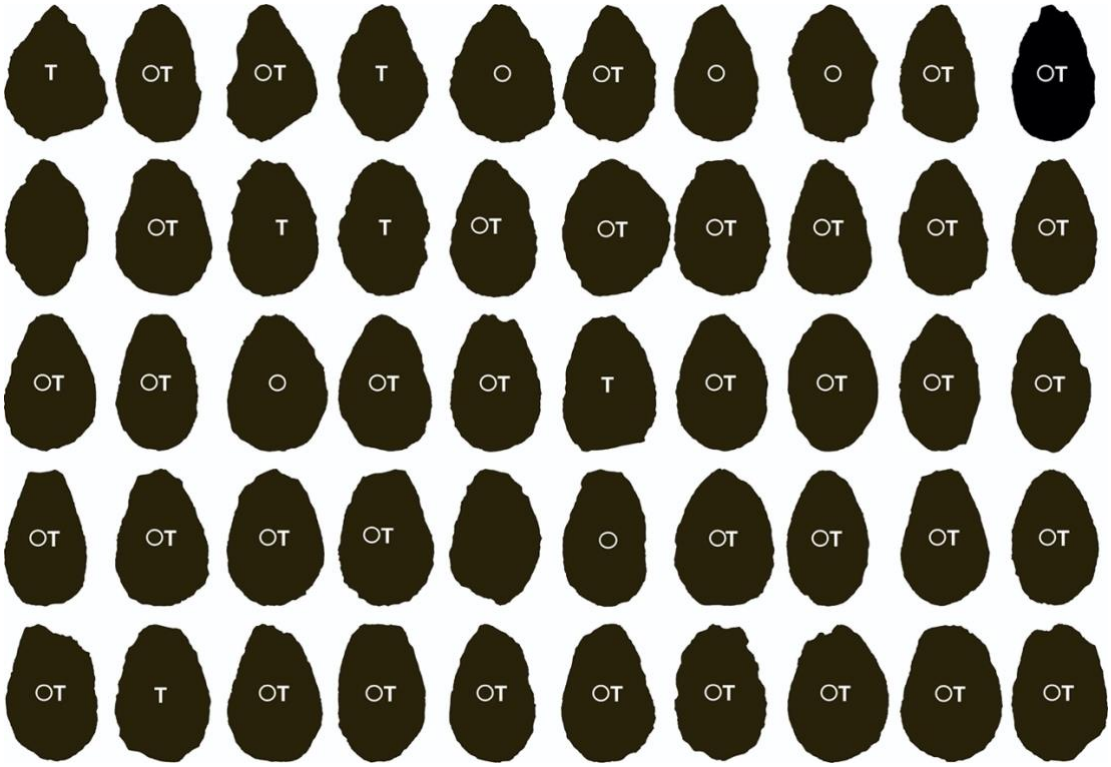


Figure 2. 6. Planform outlines of a randomized sample of handaxes from Boxgrove, Unit 4c. (Source: McNabb 2017).

More generally, Shipton and White (2020) note that at British sites, including Boxgrove, similarities within assemblages are often more marked than the variation between assemblages. They propose that this homogeneity can be explained by normativity – the conventions that exist in human social-group life and relate to the collective agreement of the group about the way people should behave (Shipton, 2010). They suggest that reproduction of similar forms over time demonstrates that makers recognise types which are linked to a socially shared understanding of the expected norms of handaxe shape.

The existence of norms amongst Acheulean groups, if true, would indicate there is meaning associated with doing something in a certain way that is negotiated by individuals within closely associated groups that share a history of learning (Giddens, 1984; Lave and Wenger, 1991; Wenger, 1998)

Therefore, clear similarities in form and process within handaxe assemblages could, as Shipton and Clarkson (2015) suggest, indicate that handaxes were made by socially connected groups who, in addition to practical necessities, shared non-utilitarian and socially significative aspects of handaxe manufacture. This could relate to group membership and identity, territory, history, and, as Barham and Everett (2021) note, potentially language.

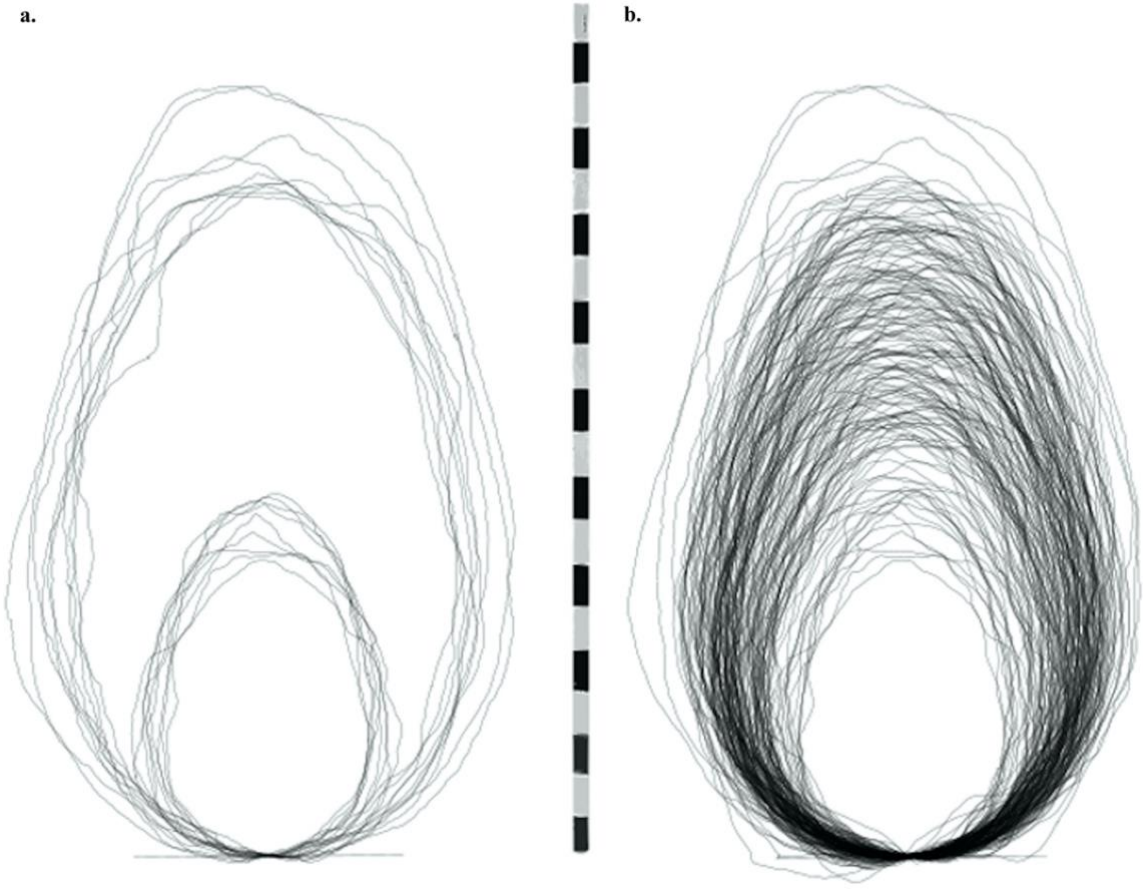


Figure 2. 7. Homogeneity of planform shape at Boxgrove. Outlines of Boxgrove handaxes drawn from Marshall Database (2002). A) 10 longest and 10 shortest. B) All handaxes. Scale in cm.

More broadly though, normativity witnessed within an assemblage can be extended to finding similarities between handaxe assemblages of broadly similar age. Utilising the British Acheulean record, Bridgland and White (2015, 2014) demonstrate similarities in handaxe shape occurring among handaxe assemblages of broadly similar age. For example, Boxgrove, High Lodge, and Warren Hill are predominantly ovate/cordiform and dated to Marine Isotope Stage (MIS) 13, c. ~524-474 kyr. Shipton and Clarkson (2015) use this to argue for the presence of temporally distinct cultural 'traditions'. Others use distinctive forms subsumed within the ovate/pointed distinction (e.g. ficrons, twisted-ovates, plano-convex and giant handaxes) to propose the existence of spatio-temporally distinct cultural templates (Wenban-Smith, 2004).

Yet, intentional, normative practice among Acheulean hominin groups is not universally accepted. As I explain in Chapter 3, the causes of handaxe morphological variation have been considerably debated. The disputes mainly deal with how much variation is really exhibited in the Acheulean, whether handaxe manufacture can be explained by genetic or social transmission, or whether there are other factors outside of the intentionality of Acheulean knappers that explain the Acheulean pattern. However, the evidence outlined in this section regarding the socio-cognitive capabilities of Acheulean hominins, the complexity of manufacture, and the necessary time and likely communication required to learn the skill of handaxe manufacture would all suggest that the development of social norms amongst knappers in the Acheulean is not too farfetched an interpretation of the homogeneity visible within certain assemblages in the Acheulean.

2.3 Summary

This brief overview of the Acheulean and handaxe technology clearly demonstrates its importance as a key landmark in human evolution. Practically, socioculturally, and cognitively, handaxe production helped shape our hominin ancestors into apex predators, thinkers, teachers, and craftspeople. Such insights help us better understand what hominin life was like, and identify similarities with ourselves, *Homo sapiens*.

Like modern humans, *Homo erectus* and *heidelbergensis* were behaviourally very variable. Terms like 'Acheulean' and 'biface' often have the effect of homogenising variability in human actions and their associated motivations (Finlayson and Warren, 2017). Albeit both terms are useful for categorising and grouping artefacts with a similar basic template, it often plays into the narratives, like Corbey *et al*.'s (2016), that 'the Acheulean' was static and void of nuance, variation or change.

Experimental understanding of handaxe manufacture, alongside ethnographic and corroboratory fossil evidence, allows us to gauge the cognitive, social, practical, and significative importance of biface manufacture by demonstrating the complexity of behaviours necessary for successful learning and skill development.

The skill of producing handaxes to the levels which are regularly visible in the archaeological record requires cognitively complex minds and socially supportive contexts which motivate learners to sustain prolonged practice. From the brief discussion on normativity above, some handaxe assemblages may demonstrate the social constraints which act upon knappers of varying experience by exhibiting marked homogeneity of form over multiple generations (i.e., Boxgrove).

Understanding that handaxe manufacture is sufficiently complex to require sustained practice within a socially motivated environment, replete with rules and expectations regarding acceptability, is pertinent to this study. It suggests hominins were responsible for the ultimate form they attempted but that they possessed variable abilities, relative to other members of the group, in achieving it.

To explore how this thesis aims to distinguish between products of individuals possessing variable skill-levels, the following chapter will explore the history of the debate surrounding handaxe morphological variation, as well as the human and non-human agencies contributing to handaxe form. I will relate this debate to the archaeological reality of skill-level variability, showing that hominin agents, via engagement with materials, are responsible for the handaxe forms they produced. When discussing hominin intentionality during manufacture, it is vital to discuss their skill, which is an intrinsic part of each knapper. What a knapper intends can, through normativity, be broadly consistent with the rest of the group and independent of experience, however, a knapper's ability to achieve their intentions is dependent on their know-how, or skill-level. Skill-level is an observable differentiator which contributes to morphological variation and can tell us more about what skill-levels were present among associated knappers at a particular site.

After presenting the morphological debate, and the relevance of skill to it, in Chapter 3, and my theoretical framework for investigating skill-level in the Acheulean in Chapter 4, I will then outline my rationale, methodology, and hypotheses. I finish this thesis by presenting my results, discussing them, and considering the wider implications of this study.

3. Chapter 3 – Handaxe morphological variation: a history of debate

3.1 Introduction

As outlined in the previous chapter, Acheulean handaxe manufacture is sufficiently complex to require a motivational social context for learning. Without a supportive social scaffold for learners, the continuation of the skill from one generation to the next is less likely due to the inherent difficulty in becoming skilled. Learners require the more experienced in the group to show them why becoming skilled in handaxe manufacture is advantageous to the survival of the group, as well as support them through frustrations and mistakes. Therefore, the process of becoming a skilled knapper is likely to contribute to visible variation in biface morphologies, even among homogenous assemblages made by groups of socioculturally associated hominins. This is due to the variable ability of knappers to execute a shared expectation of biface form.

However, much has been published (Klein, 2000; Hodgson 2009; Hosfield et al. 2018) on the 'variable sameness' (Isaac, 1977) of the Acheulean and the reasons for a perceived prolonged lack of innovation in the Lower and Middle Pleistocene records of Africa and beyond. Explanations of 'sameness' range from the existence of a shared genetic predetermination to make handaxes (Corbey et al., 2016), the cognitive deficiencies of Acheulean hominins to innovate new technologies (Shipton, 2019) to the effect demography and mobility, as well as other external influences, have had on the rate of

innovation and the need for it (Hopkinson et al., 2013). Thus, superficially, the persistence over 1.5 million years of the Acheulean handaxe could be seen as supporting a case for technological stasis. However, arguing for a static Acheulean would be to overlook the very clear variation in morphology which can be observed over time and space.

In this chapter, I provide an overview of apparent 'Acheulean conservatism', interrogate the validity of such perceptions, and summarise previous approaches to understanding handaxe morphological variability. Ultimately, I argue that intrinsic, non-human influences on biface morphology are secondary to socioculturally influenced human intentionality when explaining Acheulean morphological variability (from homogeneity at the intra-assemblage level to heterogeneity at the inter-assemblage level). I follow this by discussing how the skill-level of individual knappers is a significant and relevant aspect of the morphological variability we see in the Acheulean record – especially within relatively homogenous assemblages at the intra-site level.

3.2 Acheulean conservatism

Comparisons to post-Acheulean technological development and increased innovation rates (i.e., prepared core and blade production) can lead to the view that the Acheulean was static and homogenous across its spatiotemporal extent. This view is often accompanied by explanations that Acheulean hominins had smaller brains than later humans, as attested by the fossil cranial evidence. This in turn has been used to suggest Acheulean hominins were cognitively unable to invent and innovate in comparison to subsequent hominin species such as Neanderthals and, particularly, modern humans. Additionally, some have proposed genetic explanations for supposed morphological similarity in the Acheulean (Corbey et al., 2016). Acheulean inventiveness may indeed

have been less marked, but this is not best explained by genetics or cognitive ability, as I will discuss below. Proper consideration of Acheulean variation leads to the rejection of a homogenous and unchanging view of 1.5 million years of hominin behaviour and, instead, can reveal the significance of what 'variable' means in Glynn Isaac's famous determination of Acheulean 'variable sameness'.

Demography and mobility have been cited as factors influencing variable rates of cultural innovation. Powell et al. (2009) expand upon Henrich's (2004) model which rationalizes that increased population size determines the likelihood for greater spread and rate of cultural innovation. Powell et al. (2009) show through simulation that interconnectedness between sub-populations has the same effect on innovation and variation of artefacts due to an increased exposure to variety. However, Grove (2016) finds that encounter rate of ideas and people, not population size, is more important for determining the rate of innovation via cultural transmission, the spread of said innovation and, therefore, explaining the variation we see in the archaeological record. Grove (2016) argues that both increased population density and greater mobility are responsible for increasing the encounter rate, independently but equivalently, between people and ideas. Without a greater probability of contact, innovation distribution among populations is likely to be more localized and thus divergent in the case of non-interaction (Maier, 2015). Applying this principle to the Acheulean, Hopkinson et al. (2013) explain that lower population density, smaller group sizes and smaller territories of archaic hominins would have reduced contact and cultural transmission between groups. Therefore, the possibility that groups maintained sufficient contact to allow innovations to endure in subsequent populations would be reduced, and thus implying long-term cultural conservatism when in fact innovations may have been just as regular but more ephemeral.

The general understanding that hominins lived in small groups which, over time, gradually increased (with brains) in size (Gowlett et al., 2012) and began to innovate more, would fit the above model that Hopkinson et al. (2013) outline. Developments in stone-tool technology are broadly consistent with this model too. I discussed in Chapter 2 the innovations and changes visible in handaxe manufacture that broadly separate early from later Acheulean handaxes (e.g., greater refinement and platform preparation). This occurs alongside an increase in brain (and group) size from *Homo erectus* to *Homo heidelbergensis* and the progressive dispersal of handaxe-making hominins into Eurasia. As brain (and group) size grew further among Neanderthals and *Homo sapiens*, Prepared Core Technology innovations (PCTs), like the Levallois technique, were introduced in more densely packed landscapes that engendered increased encounter rates. Greater population density and greater mobility, leading to more numerous meetings between groups, would then give rise, following the above models, to further innovation including blade production in later Palaeolithic groups.

If Acheulean hominins experienced reduced contact between smaller, more sparsely spread-out groups, then this should have indeed reduced innovation. However, as Maier (2015) points out, this should have also simultaneously stimulated localised divergence in types of behaviour infrequently shared with other groups due to attenuated interaction and transmission of ideas. These circumstances may well explain morphological, or stylistic, differences in handaxes across time and space in the Acheulean but does not adequately explain why the bifacial template of butchery tools persisted for so long in the archaeological record.

Shipton (2019) argues that while the conservatism of form in the Acheulean is underpinned by hominin capacities for high fidelity social transmission like our own, Acheulean hominins were still cognitively unable to innovate new technologies. Normally, even amongst modern human groups, who are not unable to innovate, social (or cultural) transmission would inevitably propagate variation in the same task due to human choice, differential selection of ideas (or memes), as well as the fallibility of human replicative abilities (Gabora, 2004; Henrich et al., 2008; Lycett, 2015; Schillinger et al., 2016). Furthermore, there are influences such as demography, competition, mobility, economy, and environment, that are above the scale of the individual, or group, but influence an individual's, or group's, selection for, or modification of, an idea. Such influences on the selection and/or modification of an idea can help to explain the mutable nature of its variation and persistence. Therefore, when an idea is relatively constant and homogenous over space and time, something is stabilising the selection for it, minimising drift, and variation. Indeed, as briefly mentioned in Chapter 2, Kempe et al.'s (2012) and Schillinger et al.'s (2014) experimental results showed that size and shape characteristics of Acheulean handaxes are particularly prone to copying error and should otherwise produce mutation rates beyond what is seen in the Acheulean archaeological record (Lycett et al., 2016). Therefore, whatever the method of social intervention (imitative social learning (Schillinger et al., 2015), direct instruction (Morgan et al., 2015), or language (Stout and Chaminade, 2012)), high fidelity social transmission has been proposed to have played a role in stemming natural copy error mutation and helped to produce the relative stability of the Acheulean pattern (Lycett et al., 2016; Shipton, 2019).

Ideas, or memes, regarding culturally manufactured items can be functional (related to the use of an object) or non-functional (related to something apart from its use, such as

style). These ideas are not exclusive and coalesce in the finished object made, producing a functional item that is also styled in a certain predictable way. Persistence of such ideas relate to their 'fitness' in given sociocultural and utilitarian contexts, however, elucidating why an idea is deemed to be culturally 'fit' is harder to ascertain (Ramsey and De Block, 2015). Maier (2015:42) explains that the functional meme of a tool is constrained by the laws and conventions of its use and is 'pay-off correlated' to its individual fitness. However, Maier (2015:42) also notes that the stylistic meme of a tool is pay-off independent regarding its fitness. Therefore, functional memes exhibit less variation than stylistic ones, as style is open to increased innovation, copying error and cultural drift, because it has less impact than the function in how successfully the tool is employed (Maier, 2015:42). A contemporary example of this would be firearms. The general firing mechanism, or functional meme, is a trigger activating a spring mechanism which fires a pin into the bullet, which ignites the explosive charge in the primer which then ignites the propellant to fire the round. Although variation in shape and size of gun is linked to function (whether shooting ducks or sparrows), guns designed for similar tasks still vary greatly in style.

This explanation arguably fits the Acheulean handaxe pattern too. There is a general bifacial template (outlined in Chapter 2), but it varies spatio-temporally in shape and size. Indeed, Key and Lycett's (2017) and Key's (2019) comparative analyses of archaeological and experimental handaxe specimens demonstrates that handaxe size and planform shape can fluctuate considerably without affecting the utility of the biface. What their data shows is more tightly constrained is the profile morphology of the tool, suggesting it has clearer functional ramifications. This makes sense given the profile exhibits the working edge that forms the tool.

Therefore, although it could be argued that there is a stasis in the functional meme of the Acheulean, there is no continuity in the stylistic memes iterated in the Acheulean. The persistence of the functional meme of the Acheulean handaxe potentially vindicates Shipton's (2019) position that high-fidelity social transmission occurred amongst Acheulean hominins due to the fact the Acheulean handaxe template continued to work well over space and time. The fact that handaxe planform shape was freer to fluctuate is related to its reduced impact on the function of the tool. In this way, both the persistence of the functional meme and fluctuation over space and time of the stylistic meme of the handaxe simultaneously demonstrate the human agency behind the Acheulean pattern. The handaxe worked and as such persisted via high-fidelity transmission but its planform design was sufficiently changeable, whilst remaining functional in profile, to allow for stylistic variance that could denote cultural difference.

Indeed, stylistic variants on the bifacial template of the Acheulean handaxe have been categorized extensively and thought to represent cultural variance and preference in shape (Bordes, 1968; Bridgland and White, 2015, 2014; Roe, 1968; Shipton and Clarkson, 2015; Shipton and White, 2020; Wymer, 1968). Closer analysis by David Bridgland and Mark White (Bridgland and White, 2015, 2014), discussed above, has attempted to relatively date variations in handaxe morphology in the British context, and highlighted the non-static nature of the planform shapes over time and space. Others, such as Shipton & Clarkson (2015), argue that these spatio-temporally distinct variances in handaxe planform shape represent distinct cultural traditions.

Similarly, innovations such as the Victoria West method from 1 myr BP, which is the earliest prepared core technology known (Li et al., 2017), as well as platform preparation known from Boxgrove in the later Acheulean (Stout et al., 2014), may corroborate the views of Henrich (2004), Powell et al. (2009) and Grove (2016). Once population size and density were large enough among groups of *Homo erectus* and/or *Homo heidelbergensis* (the Acheulean was not homogenous or static in terms of who made it either), the necessary conditions were in place for slight innovations on the same concept. We also see a growing variation in lithic reduction methods during the later Acheulean that are not focused on producing bifaces (Wilkins and Chazan, 2012). These innovations were few and far between, especially in comparison to later periods in the Palaeolithic, but they are still significant in contesting arguments which assert cognitive deficiencies among hominins meant they were incapable of innovation (Shipton 2019). Certainly, technological developments, discussed in Chapter 2, such as controlled use of fire (Berna et al., 2012; Goren-Inbar et al., 2004; Gowlett et al., 2012; Wrangham, 2010) and the development of hafting and hunting weaponry (Milks et al., 2019; Thieme, 1997; Wilkins et al., 2012) would suggest Acheulean hominins, especially in the later Acheulean, were cognitively able to innovate freely as and when conditions necessitated.

Alternatively, the functional meme was mutable as well (i.e., for other tasks such as woodworking (Dominguez- Rodrigo et al., 2001)). This is a less well-explored potentiality, with most researchers assuming automatically that handaxes were butchery tools. However, examples such as the giant handaxes at Cuxton (Wenban-Smith, 2004) do not seem to be able to hold any real functional advantages. Moreover, multiple examples of oversized, as well as miniature, handaxes are difficult to visualize as having been used for processing carcasses. As such, they bring into question the universality of

the stasis of the functional meme also. Therefore, explanations, like Corbey et al.'s (2016), of handaxes as a homogenous and unchanging technological tradition, need to be challenged in light of the evidence for variation in the broad context of the Acheulean.

Corbey et al.'s (2016) principal argument for Acheulean conservatism is that handaxe form is very stable and fails to track the variation they would expect assuming different environments necessitated different solutions. Moreover, they note that the Acheulean record displays a dearth of copying errors that one would anticipate from a socially learnt skill or behaviour. Corbey et al. (2016) conclude that this suggests that handaxe manufacture was at least partly under genetic control. For them, this would explain the presence of handaxes in regions that seem unconnected, and which cover distances that might have taken more than several generations to reach.

However, Hosfield *et al.*'s (2018) and Wynn and Gowlett's (2018) responses to Corbey *et al.* (2016) have shown that a genetic explanation for the apparent Acheulean stasis is unlikely. Furthermore, a fuller appreciation of the Acheulean demonstrates quite clearly that handaxe variation does exist and conservative perceptions of it constitute a 'gross simplification of the complex archaeological record' (Hosfield et al., 2018: 10).

Hosfield et al. (2018) question Corbey et al.'s (2016) suggestion that handaxes should track environmental variation, reasoning that a hand-held butchery tool is likely to be equally functional in a variety of settings. Hence, modern knives share the same basic functional meme despite variations in their appearance. Indeed, they go on to explain that, using MacDonald's (2007) reference to hunting, it is behaviour and not material culture that more likely tracks environmental variation.

In addition, Hosfield et al. (2018) demonstrate that Corbey et al.'s (2016: 10) 'macro-scale perspective' fundamentally simplifies and misrepresents the variability of the Acheulean, and copying errors, or diversions, do exist over space and time. They (Hosfield et al. 2018) highlight something that Corbey et al. (2016) do not consider - that variability has limits that relate to concept, utility, process of manufacture and sometimes raw material. The early Stone-Age knife, or biface, was used in only a certain number of ways which limited its variability in a practical sense, but, as with knives, there was room for variety in styles. Therefore, the similarities that Corbey et al. (2016) see in the basic morphology of bifacial handaxes (i.e., that they are bifacial) is not conservative if you appreciate the full extent of the stylistic variation produced within that broad bifacial template.

Similarly, Hosfield et al. (2018) explain that the presence of regions that were predominantly free of handaxes or distant from more established handaxe producing areas is not evidence for a genetic explanation. Instead, it can be explained by mobility and population dynamics (Hosfield et al. 2018). Indeed, as well as Corbey et al.'s (2016) underestimation of the abilities of hunter-gatherers to cover substantial distances, Hosfield et al. (2018) note that, contrary to Corbey et al. (2016), regions where handaxes are absent (subject to future finds) were not necessarily maintained by a lack of suitable raw material but by other socio-cultural factors such as site function, length of occupation, and range of activities. Indeed, the fact that there are spaces and geographic areas in the Acheulean world where handaxe manufacture is not present arguably undermines a genetic hypothesis. For example, the Clactonian (a flake technology without handaxes) in Britain occurs in a geographic area where handaxes had previously been made for a

considerable period. Therefore, as Hosfield et al. (2018) note, the presence of this technology more likely indicates differences in site type or activity as there were plenty of handaxe making groups and suitable material not too far away from Clactonian sites.

Wynn and Gowlett's (2018) response similarly identifies considerable problems with the genetic explanation for handaxe manufacture (see Box 1 on page 22 in Wynn & Gowlett 2018). These focus on two related aspects of handaxe manufacture: 1) ergonomics and 2) aesthetics. They propose that in the absence of hafting technology, handaxes would provide the most efficient solution for cutting tasks and this explains the possible scenario of independent invention in widely separated regions. Secondly, they suggest that aesthetic expression is documented in handaxe forms being linked to primate pattern-detection networks and pleasure-reward systems. This can be interpreted as an aesthetic impulse that was expressed culturally. The inclination to make a handaxe was thus not genetically determined but partly reflects the desire of the maker to produce an aesthetically pleasing form on a cultural artefact. Variation in people, time, and space therefore produced divergence in aesthetic, cultural responses to comparably similar functional and ergonomic needs of hunters over the course of the Acheulean. Therefore, the same basic principle can yield considerable variability. Indeed, Wynn and Gowlett (2018) note that by the later Acheulean (after 600 kyr BP) aesthetic considerations sometimes overrode functional and ergonomic considerations for handaxe manufacture in the appearance of particularly extreme forms (e.g., giant handaxes and extreme ficrons – handaxes with long, narrow points). Their reduced utility intimates that they had some culturally significative importance beyond function. Alternatively, this could be argued to be evidence for a genetic penchant for making handaxes even when they are unable to be used, but this does not fit in well with the known picture of the Acheulean.

What is clear from these (Hosfield et al. 2018; Wynn and Gowlett, 2018) responses is that the Acheulean cannot be understood through a single explanatory lens. Although handaxes share a basic template and are predominantly bifacial, they are by no means morphologically homogenous and can vary widely in their characteristics, size, shape, and use across space and time. I would argue this shows that the variation we see in the Acheulean was intentional, motivated by the fusion of functional need and cultural history. Although it is possible that cognitive limitations may explain why other diversified technological lithic solutions for everyday butchery and survival tasks did not start to arise until the Later Acheulean, it is also possible that something other than cognition was driving the necessity for innovation in lithic technological repertoires.

Finally, even those who challenge the notion that the Acheulean was variably similar have approached handaxe morphological variation in contrasting ways. They predominantly focus on the influences of human agency versus those of external agencies in shaping the variability we see. However, not all approaches to handaxe morphological variation, have placed hominin intentionality and agency at their centre as I have argued here. Therefore, in the next section I will synthesize, briefly, a history of the main approaches to morphological variability before outlining how it is relevant to investigating skill-level differences in the Acheulean.

3.3 Principal Approaches to Morphological Variability

3.3.1 Culture Historical Approach

Proponents of the culture historical approach have viewed handaxe morphological variation from a typological perspective and as an indicator of broad group identities. They generally relied on metrical and statistical analyses to measure differences in morphology between handaxe assemblages. This has resulted in the clustering of types based on similarities in shape and in relation to their age. In doing so, culture historians moved away from earlier traditional approaches which did not concentrate overly on variation of form but identified major characteristic cultural traditions that related to specific blocks of time and space in the Palaeolithic record – i.e., handaxe equals Acheulean (de La Torre, 2016; de Mortillet, 1873; Sackett, 2014). Instead, the culture historical approach looked for variation within the Acheulean (Sackett, 2014).

Roe (1968) and Wymer (1968) working on British material and Bordes (1968) focusing on French material, are the pioneers of this approach and use of their morphological descriptions continue to this day. Their collation and cataloguing of material have produced remarkable pieces of work and the variation that they describe is still enormously valuable.

Roe (1968) sorted handaxes into two main Traditions, pointed and ovate, then sub-divided further into groups based on morphological similarity. Recently, Bridgland and White (2015, 2014) have used river terrace systems to date Roe's (1968) Groups. Wymer (1968) provided similar analyses but went into further detail and described variation in production of cutting edge, shape of edges, as well as the shape of both points and butts of handaxes. Bordes (1968) also used metrical indices of handaxes to group them according to morphological similarity. However, he distinguished between classic and

non-classic specimens and developed a different nomenclature to that of Roe and the British tradition.

Attempts to categorise handaxe variation like this were important steps in eliminating the *fossil directeur* approach (Field, 2002) and illustrated that the Acheulean was not a monolithic block or reflected technological constancy. Moreover, it highlighted the agency hominins had over deliberate shaping of their bifacial tools, not just their ability to create a handaxe.

By typologically categorising and addressing morphological variation in the records over space and time, proponents of the culture-historical method admitted to the presence of differing explanations for archaeological patterns. Put more simply, they were some of the first to try and explain what the variation might mean.

3.3.2 Utilitarian Approach

Utilitarian approaches to handaxe morphological variation attempted to explain the factors influencing shape in an alternative way. Its proponents attempted to better understand the method of making handaxes and challenged the assumed notion that handaxe morphological variation was solely the result of deliberate choices made by the knapper. They modulated the degree of agency attributed to the handaxe maker to individually shape their tools and argued for the greater agency of raw material variation and utilitarian concerns.

McPherron's (1994) re-sharpening hypothesis was a principal utilitarian approach, influenced by Dibble's (1984) earlier work testing the reduction intensity visible in Middle Palaeolithic flake scrapers. It challenged the notion that biface shape was intended. McPherron's analysis of four Acheulean sites from France and Jersey (Gouzeaucourt, Longavesnes, La Cotte de St. Brelade and Cagny-la-Garenne) sought to explain that handaxe morphological variability resulted from the natural progression of re-sharpening within the limits set by raw materials used. Focusing on tip length, the pointedness/roundness of the handaxes, and the amount of retouch visible on them from the above sites, McPherron (1994) proposed that through re-sharpening, handaxes became more ovate as they were reduced. Essentially, handaxes started off pointed and as they were used and re-sharpened, they became smaller whilst becoming more ovate in shape.

There are some problems with McPherron's (1994) hypothesis, primarily pertaining to it being proposed as a general model for handaxe morphological variation. Although it may be locally or situationally possible for such a pattern to occur, there are instances that show this model cannot be applied universally without qualification. Boxgrove, for example, yields episodes of handaxe manufacture where blocks were reduced into ovates from the very beginning and which do not exhibit intensive re-sharpening or reduction (Garcia-Medrano et al., 2019; Leroyer, 2016; Pope, 2002; Roberts and Parfitt, 1999). Iovita and McPherron (2011) later applied the re-sharpening hypothesis to Boxgrove handaxes but did not present direct evidence of re-sharpening other than shape analysis by size. Indeed, Shipton and Clarkson (2015) found that, counter to Iovita and McPherron (2011), the shape of Boxgrove handaxes, as well as handaxes from Swanscombe Middle Gravels and Broom Pits, were consciously conserved quite independently of reduction

intensity – suggesting imposition of a purposeful shape. In addition, it is unclear whether smaller ovate handaxes that exhibit greater levels of re-sharpening at sites like Gouzeaucourt were necessarily pointed (i.e., not ovate) to begin with – something I will deal with further in Chapter 5.

McPherron attributed Cagny La Garenne (CLG) and Gouzeaucourt (GZ) to the same reduction trajectory, primarily due to the more reduced handaxes from CLG resembling less reduced forms at GZ. However, this overlooked some key considerations. Firstly, CLG is dated to MIS 12/11 and GZ is dated to MIS 8, a gap of at least 74,000 years. Moreover, CLG is a mixed, palimpsest assemblage exhibiting both pointed and ovate forms of larger dimensions than bifaces found at GZ (where there is over 90% ovates) (McPherron, 1999, 1994). The fact that CLG is a palimpsest assemblage many thousands of years older than GZ is significant. This is especially true given McPherron's claim that the points and ovates from CLG were part of the same reduction trajectory. The mixed nature of palimpsest assemblages makes it just as likely that what is found at CLG is the separate expressions of handaxe traditions that are broadly dated to MIS 12/11. Therefore, finding a link, as McPherron (1999, 1994) did, between this mixed assemblage and an assemblage in MIS 8, largely unaffected by post-depositional action and yielding generally homogenous shapes, is difficult to argue in light of these considerations.

Another critique of McPherron was provided by Ashton and White (2003). They argued that due to ovates having wider midpoints than pointed forms, ovates could not have been reduced from points. In other words, if a pointed handaxe was larger to begin with but was less wide at the midpoint than a smaller ovate handaxe, then how could the midpoint grow in width whilst being reduced to an ovate? However, Emery (2010) has provided

an adequate explanation for this – that the midpoint can move its position further towards where the base of the point originally was. This is best illustrated in the below figure 3.1.

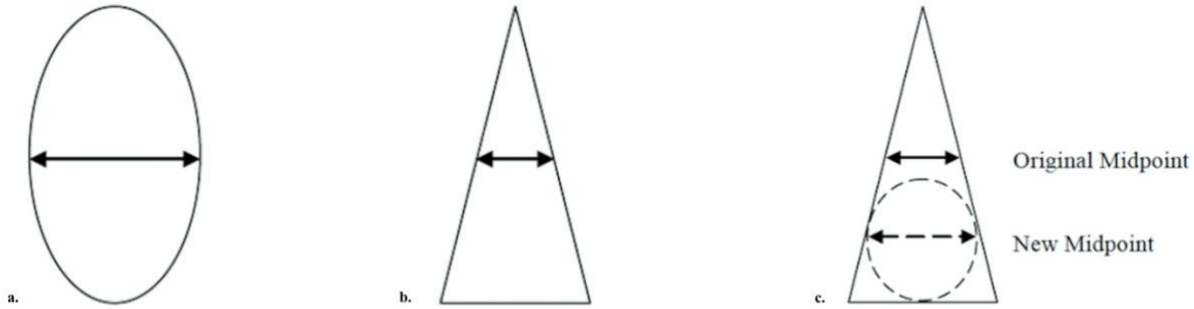


Figure 3. 1. Emery's (2010) appraisal of Ashton and White's (2003) critique of McPherron's (1994) resharpening hypothesis. Images a. and b. show that ovates have wider midpoints than points, however, image c. shows that the midpoint can move in line with reduction.

Emery's point is persuasive and makes sense. Yet, if the focus of the reduction of the handaxe is the tip, as McPherron argued, then reduction towards an ovate form would leave a rounded form that is quite thick given the fact that larger pointed forms have thicker butts. However, as McPherron (1995) identified, this is not the case and ovates are more refined than pointed forms. This suggests that if reduction did take place from pointed to ovate, as he suggested, then hominins would have had to focus on more than just reducing the tip but begin to reduce the biface more equally around the entire perimeter of the piece. Whether for ergonomic reasons or maintaining functional effectiveness, it would be more difficult to refine a biface that has been reduced to the point of becoming ovate where the reduction has been solely focused on the tip.

If, as the data suggests, smaller ovates for whatever reason needed to be more refined than larger pointed handaxes, it would have made more sense to reduce the entire piece

more equally with this in mind. It would have allowed more material, and therefore more possibilities, for the necessary long thinning flakes needed to refine a biface. If what was the thicker butt of the point remained static whilst the tip was the focus of reduction, then arguably it would be more difficult to produce the smaller refined handaxes that McPherron argued were the end result of this resharpening process. This is because it would necessitate a change in re-sharpening approach and less material for which to remove the thickness of the piece.

Like McPherron (1994), Davidson and Noble's (1993, 1989) concept of the 'finished artefact fallacy' challenged the notion that human intention was the principal source of biface morphological variation. They proposed that the form of a lithic tool, when discovered, only represents one point along a continuum of potentially multiple modifications (Dibble et al., 2017). It was used to suggest that a 'desired end product' is incorrect and something that has been imposed upon Acheulean handaxe makers by archaeologists.

Although I do not necessarily agree with the presented narratives, I concede that the artefacts found in the archaeological record do have the potential to be re-sharpened and to change their form. Moreover, the same form could be used for different purposes (i.e., cutting or scraping). However, this fact does not preclude a tool being deliberately shaped. All it excludes is determining correctly what previous deliberate shaping episodes took place earlier in an artefact's 'life'. Also, the form in which we find handaxes represent, by definition, the finished form of the artefact, even if they were previously re-sharpened and were intended to be re-sharpened later but were not. Previous iterations along a re-

sharpening chain do not disqualify the notion that, in each instance, whoever was resharpening a biface, had a particular goal in mind.

Indeed, Ashton and McNabb (1994) and White's (1998) raw material hypotheses imbued the Acheulean knapper with more agency, arguing that ovate morphologies were usually the desired form when feasible (Ashton and White, 2003; Emery, 2010). However, they proposed that raw material type, size and quality was the principal determining factor affecting handaxe morphology. White's (1998) study of multiple British handaxe assemblages maintained that, in essence, pointed forms were the unavoidable result of using poor quality, often river transported and elongated raw material. He argued that poor quality material conditioned the actions of the knapper to make the easiest shape possible, thereby precluding comprehensive knapping of more than one end of the tool (Emery, 2010; White, 1998). Consequently, White considered the presence of ovates as evidence of a lack of external constraint on the knapper and that they could create whatever they wanted. To White, this related to a generalised mental template consisting of ovates with continuous sharp edges and a high level of symmetry (Emery, 2010; White, 1998). However, he did accept that there is a high degree of variation within shape groups but maintained that points and ovates are distinctive shape-classes.

I do not disagree with this interpretation (of points and ovates being distinctive or that raw material is an important determining factor in the morphology of the handaxe), but I would challenge the notion that there is a generalised template among handaxe makers. White (*ibid.*) considered a multitude of variations in his analysis, yet he subsumed more atypical morphologies, such as ficron, plano-convex, twisted ovates and others within the main ovate/pointed classes.

An alternative to the utilitarian hypotheses outlined above was put forward by Wenban-Smith (2004). He used the giant ficron found at Cuxton to argue in favour of deliberate, non-practical motives and socio-cultural agencies sometimes affecting handaxe morphology. However, it remains to be demonstrated whether poor quality raw material really does hinder a knapper's ability to produce ovate handaxes. I do not necessarily concur with the view that making pointed forms on poor quality material is to follow a 'path of least resistance' (Ashton and McNabb, 1994; Emery, 2010; White, 1998). Instead, if there is a preferred form and hypothetically no time constraints involved, would they make a pointed form just because it was easier? Not necessarily. Eren *et al.*'s (2014) experimental study has shown that difference in raw material is not a predominant factor in determining handaxe shape (Lycett et al., 2016). For me, therefore, both raw material and re-sharpening hypotheses neglect the agency of non-utilitarian hominin intentions.

Utilitarian and non-utilitarian influences on handaxe morphology are knitted together well in the studies by Key (2019) and Key and Lycett (2017), discussed above. Their research examines the relationship between utilitarian considerations of effective functional use of the handaxe for butchery purposes and biface morphology (planform and profile).

Key and Lycett (2017) found that handaxe shape variation can occur independently of cutting efficiency, although there are task specific thresholds below a certain mass of tool. Additionally, Key's (2019) functional analysis of large archaeological and experimental samples of handaxes concur with the notion that handaxe planview shape is more

reflective of cultural tradition and not raw material agency or the unintentional consequences of reduction. He suggests that the comparatively increased planview shape variability among archaeological specimens demonstrates that this aspect of handaxes may have been related to other factors than utilitarian issues of effective handaxe use. He suggests that handaxe planview shape may have been subject to cultural drift within hominin systems of social transmission. However, profile morphology among archaeological specimens has a much tighter range than their planview shape. Key (2019) argues this represents utilitarian constraints on the use of handaxes. He relates this to the amount of load that can be exerted on them during butchery/cutting tasks, as well as the ease with which they can be gripped.

This analysis would suggest that specific planform templates can be imposed by hominins regardless of what material they are using. More interestingly, reduced profile morphological variation among archaeological handaxes compared to experimentally made ones implies that utilitarian constraints existed. This utilitarian outlook differs from the other utilitarian hypotheses outlined above. It does not attribute variation in morphology as the result of external or unintentional processes. Instead, it assesses biface manufacture and use, the role of the material and what it can withstand for its supposed purpose and relates this back to biface morphological variation. In so doing, Key and Lycett (2017) and Key (2019) have returned agency to the hominin, showing how understanding of what is useful/not useful for their tool constrained the functional meme (profile morphology) and allowed the stylistic meme (planform morphology) to drift in line with sociocultural influence. Both aspects delineate intentionality behind biface morphological variation.

3.3.3 Holistic Approach

Monocausal, or predominant factor, explanatory models of handaxe morphological variation, like McPherron's outlined above, attempt to comprehensively explain and rationalise Acheulean variability in its entirety. However, they fail to properly consider other possible agencies of variability, as well as contextual specificities potentially relevant to interpreting differences. Instead, a more holistic approach offers better possibilities for addressing the question of variation.

Anna Machin's (2009) multi-factorial model offers such an approach. This model explains variability in handaxe morphology by concentrating on the individual agent producing handaxes at any given point in time. The model predicts that variability is caused by a multitude of differing social motivations and ecological, physiological, biological, and cognitive constraints, which is summed up quite chaotically in Figure 3.2 (Machin, 2009). However, Figure 3.2 also adroitly illustrates the contextually specific and unique series of events that relate to each individual handaxe produced. Furthermore, Machin's treatment of potential socio-cultural agents of handaxe variability, such as social group dynamics, identity, specific manufacture method, physiology, and cognitive constraining factors, help provide a more balanced view of how and why dissimilarities between handaxes exist. It also suggests that analysis of the individual through stone-tools is viable if one understands the interplay of agencies between individuals and their surroundings. This is because the individual is the driving force in making the handaxe. A holistic approach, therefore, can investigate patterns of variation among handaxes that have been produced by the same group and not just broadscale patterns of variation.

Investigations of the ‘individual’ is a concept thought un-promising for the earlier periods of the Palaeolithic, with some deeming such a ‘bottom up’ approach as tantamount to ‘theoretical storytelling’ (Foulds, 2014; Gamble and Porr, 2005). However, as handaxes are the products of individuals, there should be something to say about the individual who made them. This is sufficient cause for focusing on identifying aspects of the individual knapper that are unique to them and only discernible through analysis of Acheulean handaxes.

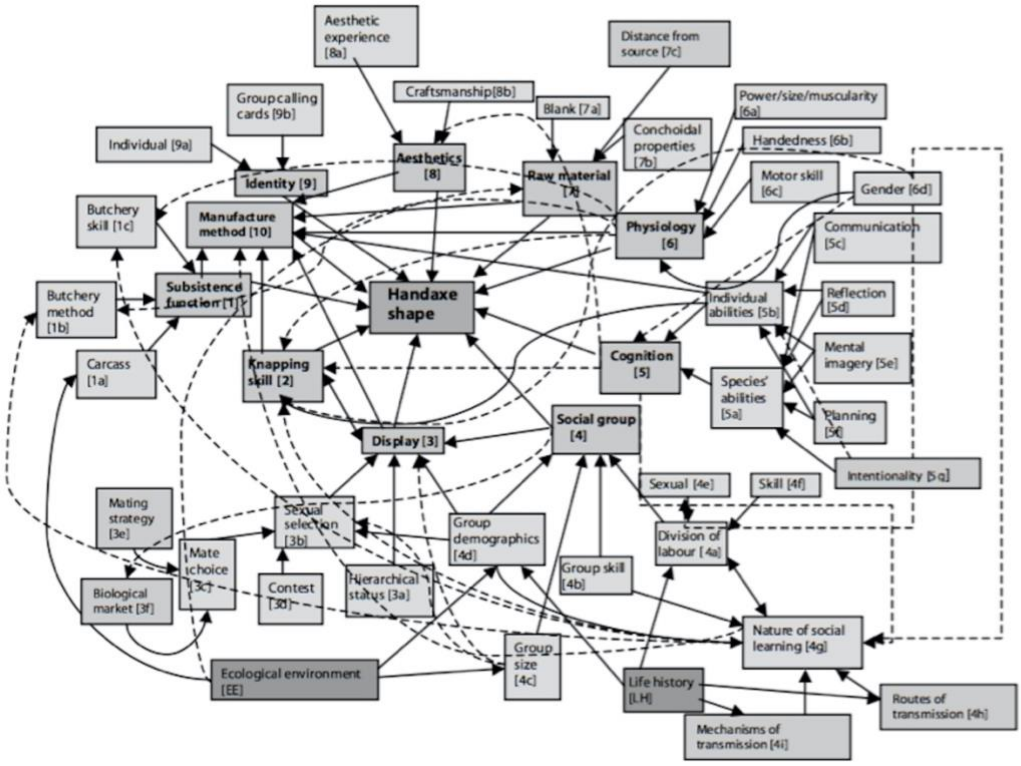


Figure 3. 2. All factors that influence handaxe shape. (Source: Machin 2009).

One aspect of the individual knapper that is potentially knowable and highly variable is their skill-level. Here I must re-iterate that it takes considerable practice and effort to become skilled in Acheulean handaxe manufacture. Therefore, at any one time in a group,

or in the wider landscape, there would be knappers of varying abilities contributing to the archaeological record (Bamforth and Finlay, 2008; Finlay, 2015, 2011; Shea, 2006). This would be the case whether the whole group participated in learning handaxe manufacture, or the practice was consigned to a select few of 'specialists'. As such, a determination of skill-level and its relation to handaxe morphological variation through analysis of Acheulean handaxes has the potential to provide information about the individual through comparison with others within the wider group. As such, it is not a pure examination or investigation of the individual but of categories of them in context within the wider group. Such an approach has been developed by Leroyer (2016).

Like Machin (2009), Leroyer (2016) takes issue with overly utilitarian explanations of handaxe manufacture and considers socio-economic factors more closely. However, he does not ignore them completely. Instead, Leroyer advocates for a holistic (or systemic) approach to the analysis and explanation of handaxes and their patterns. He explains that considering the social and individual factors that influence the handaxes produced as part of a holistic system, where functional, utilitarian, raw material and other considerations are similarly considered, is crucial to understanding everything relevant that contributes to the context of handaxe manufacture. Leroyer's methodology is based on a very inductive postulate: that Acheulean handaxes are the result of complex processes.

Leroyer analyses the operating chains of handaxe production to contextually understand the multiple factors (including individual/human, raw material quality and shape, functional objectives, cultural norms, and the life span of the tools) and their interaction, or engagement, with each other. He explains that the analysis of operating chains should not prejudge *a priori* the influence of one factor (e.g., individual-related variations) and

should not attempt to demonstrate the link between the observed variations and this single factor independently of any other.

Leroyer's analysis leads him to conclude that social, or cultural, norms related to what was deemed acceptable manufacture were affecting the knappers at Boxgrove to a point where it was visible despite raw material differences. This is supported by Garcia-Medrano et al.'s (2019) analysis of the large assemblage of handaxes from area Q1/B (*n* = 358) at Boxgrove. Their analysis shows that raw material size and shape differences and highly variable reduction sequences produced heterogenous handaxe sizes but homogenous ultimate shapes (i.e., a 'common goal' was attempted by Boxgrove knappers). This is further corroborated by Shipton and Clarkson's (2015) analysis demonstrating that shape was conserved at Boxgrove regardless of reduction intensity.

3.3.4 Summary

Most of the debates surrounding interpretations of the causes of handaxe morphological variability centre around agency and perceptions of the abilities of hominins. Monocausal explanations of handaxe morphological variation, as well as genetic explanations for apparent 'stasis', are insufficient in explaining localised and/or broader patterns. It appears equally unlikely that the variation visible in the Acheulean handaxe record was the result of differences in raw material alone and that the reality of handaxe manufacture across the spatio-temporal extent of the Acheulean is far more complex and exists in the engagement of human and non-human agencies in a social context.

Engagement of human and non-human agencies is the basis for understanding variation at multiple scales of analysis. It allows for the acknowledgement that despite raw materials affecting handaxe manufacture in each instance there are human motivations involved in the production of certain handaxe shapes. Machin's (2009) holistic approach highlights the reality of one's engagement with a material. It is in, as well as out of, their control and is not static. An individual's ability in working with variable materials is their own and, as such, is within their control and improves with increased practice. One cannot control the variability of material encountered on a daily basis. However, with increased exposure to, and practice with, heterogenous lithic materials, one is more in control of one's engagement with the raw material in general through knowing the ways in which a material can behave. This can only happen through sustained practice and success in overcoming problems encountered.

One's ability, then, must be a crucial aspect of the visible variation in handaxes in the Acheulean record. This, as will be discussed below, is rarely acknowledged and a potential avenue for gleaning more from the inert stone-tools we are left with. A considerable challenge is going beyond acknowledging what hominins did, and to understand, in some way, aspects of who they were. I do not mean 'who they were' in terms of how they thought of themselves, their beliefs, or specifics about their person but, instead, who was present? The skill-level of an individual is inherent to them and as such should be a knowable thing about them in relation to their wider group. Therefore, can investigations into the skill-levels of individuals contributing to the Acheulean record yield a higher-resolution image of the Acheulean?

3.4 Relevance of skill-level to debate

My initial interest was related to what skill 'is'. Consequently, I investigated what approaches have been devised to understand skill more generally - from the perspectives of both the individual and observer. I will touch on these approaches more in the theoretical framework outlined in the following chapter (4), but they reveal that the skill-level of an individual is necessarily personal and embodied. However, skill-levels can be compared when two or more individuals share an objective.

Applying this logic to Acheulean morphological variation, investigating skill-level can be used as a method to distinguish between individuals at a site like Boxgrove, where handaxes were made to a shared template and were in close association to each other. This is because skill-level has a significant effect on one's ability to adhere to a shared template. Therefore, because handaxe manufacture takes considerable effort and practice to become skilled, skill-level differences should be visible between two associated handaxes made by differently skilled individuals (Bamforth and Finlay, 2008; Finlay, 2011, 2015; Shea, 2006). Indeed, those also interested in skill level in the Acheulean work on the premise that there is a continuum of skill, from 'Novice' to 'Expert' (Pargeter et al., 2019; Shipton and Nielsen, 2018; Stout and Khreisheh, 2015). However, the continuum of skill cannot be universally applied across space and time. Skill is relative to what specific task is being attempted by members of a particular group and their proficiency in performing it.

This approach considers skill as relational between individuals who adhere to a collective repertoire of practice which may be part of a shared history of learning. Investigations of

skill are less relevant if one compares two spatiotemporally distant assemblages that represent groups of individuals who are creating differently shaped/styled handaxes in different ways.

Although Shea (2020) warns against treating similarities among stone tools as more important than differences, determining a shared goal can act as a helpful control variable with which to assess differences in skill-level. This is because it permits the comparative assessment of how successfully a knapper has performed within the parameters of a morphologically similar pursuit. Consequently, in support of Shea's (2020) view that differences are more important than similarities, investigations into skill-level among a relatively morphologically homogenous assemblage can potentially provide useful insights regarding the differences that exist between bifaces at the intra-assemblage level.

For example, demonstrating a common morphological goal or manufactural standard at Boxgrove allowed Leroyer (2016) to investigate categories of skill at the site, once the whole range of potential factors affecting handaxes in a localised context were considered. Although skill is personal to the individual, Leroyer was right in reasoning that it is very difficult to recognise the products of discrete individuals due to the multiple parameters that influence the shaping process, as Foulds (2014) has shown elsewhere. Therefore, Leroyer found it unnecessary to identify individuals to demonstrate that there are indeed different levels of know-how exhibited at Boxgrove. This can be applied to the archaeological record more generally. Not only is it rare to find the requisite resolution for individuals to be identified, such as at Pincevent (Ploux, 1989), Etiolles (Pigeot, 1990), and Maastricht-Belvédère (Stapert, 2007), but it is not necessary for generating

meaningful sociological interpretations of handaxe assemblages. Leroyer (2016) has shown that determinations of relative skill-level can yield equally important insights.

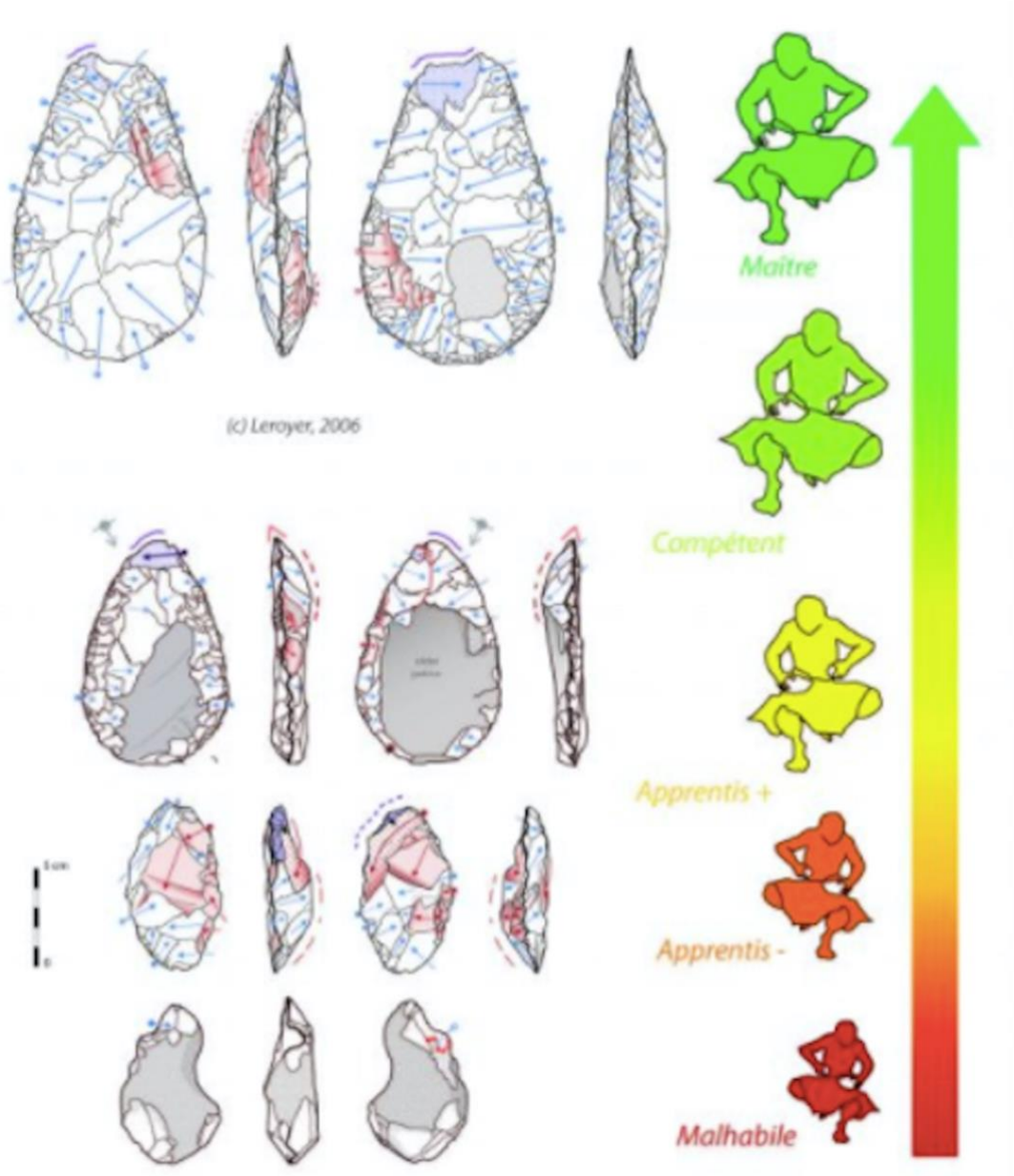


Figure 3. 3. Leroyer's (2016) skill-spectrum at Boxgrove. (Source: Leroyer 2016).

Leroyer (2016) showed a pre-determination at the start of the operating chain of manufacture towards tranchet blows that, in conjunction with the general morphological

similarity at Boxgrove, is helpful to corroborate the closeness of association of the bifaces and the existence of an 'acceptable' result of manufacture. He distinguished between five categories of skill-level at Boxgrove (1. *Malhabile* (Clumsy), 2. *Apprentis* - (inexperienced apprentice), 3. *Apprentis* + (experienced apprentice), 4. *Competent* (Skilled), and 5. *Maitre* (Master)) (see Figure 3.3). He used these to scrutinize the social dynamics involved in the creation of handaxe assemblages at Boxgrove. Leroyer's analysis of the spatial distribution of these skill categories demonstrated a spatial segregation between more competent knappers at single-use activity sites and less skillful knappers at certain 'communal' points around bodies of water. Therefore, determination of skill-level has the potential to provide information regarding who was involved in certain activities, especially if one considers skill-level to be correlated with age and/or status within a group. The pattern from Boxgrove would suggest that only individuals of a certain skill-level generally participated in hunts, something that could only have been hypothesized or theorised before such work.

Morphological variation is thereby relevant to skill-level and vice versa. Moreover, determination of skill-level categories within a group assemblage is a knowable thing, albeit in highly localised settings. Assessment of skill-level categories can potentially identify activities undertaken in different locales and promote discussion of intra-group hominin social dynamics. Efforts should be expanded to investigate skill-level of knappers in the Acheulean. This is what I intend to do in this thesis.

4. Chapter 4 – Theoretical framework

4.1 Definitions of and approaches to understanding skill in the Acheulean

The term ‘skill’, in the context of the Acheulean, has been used by some (Bordaz, 1971; Bordes, 1971; Callahan, 1979; Edwards, 2001; Muller et al., 2017; Schick and Toth, 1993; Shipton, 2018; Shipton and Nielsen, 2018) to highlight the developments of handaxe technology over time. They show that handaxe manufacture in the later Acheulean would have necessitated greater investment of time and practice to learn to achieve skilful results. Indeed, the notional split between Early and Late Acheulean largely rests on the refinement, greater symmetry, stylistic variation, and complexity of the handaxes produced. However, I move away from this definition of skill, that relates more to technological development over time, and focus more on the embodied, personal experiences and knowledge that constitute skill in specific, contextualised scenarios.

For example, the innovation of platform preparation, permitting the removal of more extensive thinning flakes, in the later Acheulean can be used to illustrate the differences in the definitions of a more generalised technological skill versus a more personal and embodied skill (Shipton, 2018; Stout et al., 2014). Platform preparation increases the number of operational processes performed. This can be interpreted as raising the difficulty of the handaxe manufacturing process. In turn, it could suggest that a handaxe made utilising platform preparation requires greater skill to make than one that does not utilise platform preparation in its manufacture, due to this increased difficulty.

In one way, this is correct as it requires the performance of a greater number of processes. However, platform preparation being a more complicated manufacturing method is not evidence that those hominins who made handaxes prior to its innovation were any less skilful and would be unable to accomplish it. The reality is that groups (pre-innovation) fundamentally did not have it as part of their skilful repertoire. By this logic, pre-innovation groups would be judged, unfairly, to be less skilful.

This technological definition of skill determines differences between groups and not between individuals within groups. Assessment of differences in manufacture between handaxe assemblages can track spatio-temporal developments in cognition, technological complexity, and behavioural variation. However, revealing anything about the range of skill-levels among groups of handaxe-making hominins, a key main objective of this thesis, can only be achieved through examination of variable performance within a single assemblage of consistently made bifaces.

My definition of skill, therefore, considers skill as something to be understood relationally between associated practitioners of a craft in space and time, as I outlined in the final section of Chapter 3. One's own skill is embodied and understood personally, yet it is also comprehended in relation to normative understandings of acceptability. Within a related group of practitioners, any individual can compare their ability to anothers'. Therefore, I have chosen to compare bifaces from Gouzeaucourt Level G (GZnivG) because the assemblage exhibits a marked techno-functional and morphological similarity (Field, 2005; McPherron, 1999; Soriano, 2000).

Understanding the variability in performance between handaxes that share a uniform morphology and manufacturing method (i.e., common practice) should generate data regarding which skill-levels (Novice, Intermediate, Expert) contributed to the assemblage (Bamforth and Finlay, 2008; Finlay, 2015, 2011; Shea, 2006). The possible insights this could yield, regarding the social make-up of specific Acheulean groups, is a more relevant method of investigating differences in skill-level between bifaces than comparing broad technological differences across biface assemblages.

The following sections provide an overview of the theoretical archaeological, neurophysiological, ethnographic, and experimental archaeological approaches to understanding skill that have furnished my methodological approach to investigating skill-level differences that will be outlined in Chapter 6.

4.2 Approaches to understanding skill

4.2.1 Knowledge and know-how

To know how to successfully do something requires knowing what to do, but to know what to do is insufficient to be able to know how to do something. Herein lies the fundamental distinction between those who are unskilled and *know of* and those who are skillful and *know how* - so eloquently distinguished into *connaissance* (theoretical knowledge) and *savoir-faire* (know-how) by Jacques Pelegrin (1990). This is sometimes referred to as procedural knowledge (knowing how to do something) and declarative

knowledge (knowing about something) (Rose, 2004). If one possesses know-how, or procedural knowledge, they know what to do and how to do it.

But how does one know what to do and then set about being able to do it? This question is particularly pertinent when discussing the more complex skills, like handaxe manufacture, performed by our hominin ancestors. The production of handaxes required hominins to acknowledge and understand multiple, and often hierarchically nested, actions and/or principles necessary to the successful completion of the task (Stout, 2011). They attained this understanding through observation, imitation, and/or pedagogical instruction, and then sustained practice. However, knowing how to perform these actions skillfully was not *transmitted* with high fidelity from more highly skilled individuals to novices. One's knowledge of how to do something must have been developed through learning and practice and, as Tim Ingold was advised by the Saami of north-eastern Finland, to 'know for yourself!' (Ingold, 2013:1). Indeed, Ingold (2013) conveys the importance of practice for one's understanding of a skill or task. He elaborates by explaining that provision of information by one, more skilled, individual to another, less skilled, individual 'holds no guarantee of knowledge, let alone of understanding' (Ingold, 2013:1 & 9). As questionnaire participant 1 notes, 'archaeologists have over emphasized the abstract 'knowledge' component of handaxe production (especially in early-learning stages) at the expense of basic motor control' (see Catalogue 2, response to Q.18, pg.17). Therefore, hominin knappers must have learnt for themselves and needed motivation to continue to practice and develop in the face of mistakes (Botha, 2015; Högberg et al., 2015; Muller et al., 2017; Shipton and Nielsen, 2018; Terashima, 2013; Whiten, 2015).

Failure in the demonstration of these hierarchically nested skilled actions usually can be visibly displayed in the final product and related to the know-how of the maker. In this respect, the performance of individual actions is part of the success of the final skilled behaviour performed and/or ultimate object created. Identifying the ability to which an individual performs these shared actions is possible using the quantitative and qualitative data outlined in my methodology (Chapter 6).

4.2.2 Traditional internalist versus Material Engagement Theory (MET)

A traditional cognitive, or internalist, interpretation of the marked techno-functional and morphological similarity of bifaces within an assemblage like GZnivG or Boxgrove would profess the existence of an internalised mental template that is shared by the knappers and intentionally imposed externally in the raw material (Gowlett, 1984; Pelegrin, 2005, 1993; Wynn, 2002, 1995).

Viewing artefact production in this way is argued by some (Ingold, 2013, 2000; Malafouris, 2013, 2010) to be hylomorphic (giving form to the material) and too rigid, neglecting the agency of the material substrate and other materials in the production of an artefact. However, internalists do still appreciate the agency of materials to affect a knapper's actions. Yet they maintain that the principal agency in shaping the handaxe is attributed to the knapper's ability to circumvent material hurdles and produce the mental image desired. Indeed, Garcia-Medrano et al. (2019) note that, at Boxgrove, the changing of hammers (from hard to soft) towards the thinning stage of handaxe manufacture was important (Roberts and Parfitt, 1999; Stout et al., 2014; Wenban-Smith, 1989). Moreover,

Pelegrin (1993) notes that any specific knapping process is not reproducible. This is due to the variability of the raw material which necessitates constant correction of the state of the stone by the knapper to achieve the mental images they intend to reproduce.

However, the internalist approach does not sufficiently consider the importance of practice in relation to the development of skill and one's ability to realise the mental template they may possess. As mentioned above, knowing what to do is insufficient for knowing how to do something. A mental template presupposes that a knapper knows what to do in relation to functional and socially acceptable standards of action. Yet, it also assumes that possessing a conceptual model of what to do endows the knapper with knowledge of how to successfully realise their internalised plans of action. Indeed, Garcia-Medrano et al. (2019) argue that 'pathways to production' were 'set in the knapper's mind' before manufacture. However, if this is the case, how is this content acquired by the individual if one cannot transmit know-how because it must be developed through individual practice?

This suggests that a mental template relates more to a generalised idea of what an individual is trying to do and does not adequately communicate in what way they know how to skillfully achieve what they are attempting. Consequently, a conceptual framework consisting of social rules and norms regarding acceptable manufacture can be transferred between individuals of the same group, laying the foundation for more homogenous forms. However, learning to successfully reproduce such a template is achieved through repeatedly engaging with the material and overall task. Knappers do not adhere to a commonly shared goal by receiving a preformed template with preconceived 'pathways to production' to perform.

Proponents of Material Engagement Theory (MET) first proffered this view. They do not consider intentions, regardless of skill-level, to be preconceived in the mind and realised in material (Malafouris, 2010). Instead, handaxes are judged to emerge from the engagement of human intentionality with the possibilities a material 'affords' (Gibson, 1979) during manufacture (Knappett, 2004; Malafouris, 2010). Intentions and affordances are not inherent properties of people and things but are relational products of material engagement in the specific knapping action. Engagement 'brings forth' (Malafouris, 2013: 173) the knapper's intention and is constituted by material 'affordances' (Gibson, 1979). This has important implications for how we think about skill.

It is through regular engagement with the material that skills and embodied understanding of sensorimotor contingency, in relation to bodily movements, develop. For Material Engagement Theorists this highlights that one's practical know-how is developed through experience of the task in the outside world. This experience acts as external memory, when the task is revisited, of the sensory consequences of particular behaviours (Iliopoulos, 2019). Therefore, know-how does not require the utilisation of an internal representation (Degenaar and O'Regan, 2017; O'Regan and Noë, 2001).

As Malafouris (2013) explains, the best angles for flake removal are not imagined by the knapper but 'discovered in action' in relation to the behaviour of the material. Hence, specific actions are not pre-planned but are at least partly constituted by the material itself (Malafouris, 2013). This explains why one can know what to do but not know how to do it. Individuals need to experience action and develop recall mechanisms regarding how

to operate successfully in various scenarios. This explains why METs consider a preconceived mental template to have little bearing on the handaxe produced.

Bruineberg and Rietveld (2014) and Rietveld and Kiverstein (2014) go further by highlighting the impact of social constraints on the emergence of intentions, suggesting that social influences affect what affordances are deemed relevant by the knapper in action. Traditionally, affordances are considered to be the possibilities for action that an environment offers an animal (Bruineberg and Rietveld, 2014; Gibson, 1979; Rietveld and Kiverstein, 2014). However, Bruineberg and Rietveld (2014) and Rietveld and Kiverstein (2014) explain that the affordances offered to an organism by the environment also depend on the abilities, or skills, the animal possesses as part of their 'Form of Life' – a development of Gibson's (1979) ideas that links relevance to the ecological niche of the animal. Essentially, an organism's environment and their abilities in relation to their material environment will dictate what is perceived as relevant in a specific situation. They explain, however, that the human 'form of life', which can be extended to Palaeolithic hominins, is sociocultural in nature which adds an additional caveat to the affordances perceptible in the material environment. The hominin environment, therefore, becomes 'sociomaterial' (Bruineberg and Rietveld, 2014; Rietveld and Kiverstein, 2014). This signifies that human/hominin engagement with affordances moves from possibilities for action to possibilities for acceptable action given societal practices which, by their very nature, have constraints on how an individual can behave.

Rietveld and Kiverstein (2014) expand on this, explaining that normative cultural practices affect an agent's perception of, or ability to distinguish, correct from incorrect, optimal from suboptimal, and acceptable from unacceptable affordances in action. They

describe this as *situated normativity*, where adequacy depends, in part, on agreement with other members of a sociocultural practice. Lave and Wenger (1991) have also noted that becoming skilled is meaningful and related to developing one's social identity as a practitioner within a 'community of practice'. An 'agreement' on the acceptable boundaries of competent practice is developed as part of a shared history of learning where group members engage with each other, negotiate the meaning and direction of their joint enterprise, utilise a shared repertoire of actions, and take responsibility for the realisation of the stated goal through 'regimes of accountability'.

In an Acheulean context, Shipton and White (2020) suggest that to reproduce similar forms over extended periods of time, makers would have had to recognise them as particular types which were linked to a socially shared understanding of expected norms of handaxe shape. The notion of a 'right way' to act in a certain practice brings forth the notion of intentionality in relation to perceiving affordances. It suggests that affordances are notably curtailed in the human 'form of life'.

However, against this, Malafouris (2020a) argues that social influences constraining the emergence of action do not have to be consciously adhered to on behalf of the agent. He describes the ability to distinguish what is relevant out of the range of acceptable affordances as 'skilled intentionality'. 'Skilled intentionality' results from sufficient learning and material engagement in a landscape of affordances that is influenced by normative standards (Bruineberg and Rietveld, 2014; Malafouris, 2020a; Rietveld and Kiverstein, 2014). Bruineberg and Rietveld (2014) explain that an agent with skilled intentions maintains a grip on multiple (acceptable) affordances simultaneously and can determine in action the optimal movements to take in a particular context. However, their

grip on these affordances is not reflective but, in line with the concepts of material engagement, is developed through repeated experience of the sensorimotor consequences of action in a situated normative sociocultural context (Bruineberg and Rietveld, 2014; Malafouris, 2020a; Rietveld and Kiverstein, 2014). Malafouris (2020a) uses this ecological approach, specifically in relation to handaxe manufacture, to provide an account of how there can be intentionality without conscious mental content for 'selection'. He disagrees that normative standards require the sharing of 'mental templates' and argues that they can be maintained over time as part of groups' skilled activity (Malafouris, 2020a).

Alternatively, it is possible that the approaches of both internalists and METs can be used harmoniously. Hutchence and Scott (2021) argue that generalised intentions, relating to the normative standards of a group and constituting a mental template of what ideally a knapper aims to accomplish, exist before and during manufacture. At the same time, specific intentions relating to individual knapping actions emerge during performance and cannot be preplanned using a mental template. Yet, the intentions that emerge during manufacture can be constrained by the mental template that exists before and during manufacture and help to restrict the emergence of the ultimate biface shape in line with the common goal. Similar ideas are held by some of the modern knappers who contributed to my questionnaire (see Catalogue 2, responses to Q.14a and 14b, pgs. 13-14).

Hutchence and Scott (2021) thus argue that the Boxgrove template of ovate handaxes with tranchet removals is evidence that mental templates and emergent intentions act concurrently during handaxe manufacture. Tranchet sharpening at Boxgrove was common and Leroyer (2016) even demonstrates that Boxgrove knappers employed a pre-

determined method for achieving it. Therefore, it could be argued that here tranchet sharpening formed part of a generalised model of handaxe manufacture. However, as can be seen in Figure 2.5 (Chapter 2), the specific tranchet action was not pre-planned in a content-specific way in the mind of the knapper. The difference in tranchet sizes, shapes, and extent show that Malafouris is undoubtedly correct in saying that the tranchet emerged as possible when the engagement of intentions and affordances permitted its removal. However, the ubiquity of the action, in conjunction with the marked morphological similarity at Boxgrove, suggests that counter to Malafouris, the emergence of the tranchet referred to a consciously held, reflective intention.

The relevance of this debate in regard to the handaxes from GZnivG relates to assessing the ability of the knapper to constrain their emergent intentions in relation to the shared common goal that existed at Gouzeaucourt. This common goal is demonstrated via shape analysis in the following chapter (5). It highlights the suitability of utilising the biface assemblage from GZnivG to analyse the relative skill-level differences between associated knappers with similar aims. Furthermore, quantitative and qualitative analyses of the outcomes of knapper actions exhibited by the end product will assess the grip to which a knapper had control of the emergent affordances in relation to the relevant actions that were necessary for them to perform in line with this socially shared template. Poor performance, inability to overcome mistakes, or failure to adhere to aspects of a template will provide comparative data related to the skill-level of the knapper.

Using the approaches outlined above, and explained in more detail in Chapter 6, I shall be able to make a number of predictions relating to both experimental and archaeological investigations of skill. For example, regarding the handaxes from GZnivG I should be

able to discern at least two out of three skill-level categories (Novice, Intermediate, and Expert). Inevitably, there is bound to be some misclassification between adjacent skill-level categories when quantitative and qualitative assessments are compared. Moreover, I also predict that because Novices and Experts constitute the extremes of the skill spectrum, they will be rarer than the Intermediates category at GZnivG. Nevertheless, given the far greater grip over affordances and material engagement that Experts supposedly possess, it is unlikely that any Gouzeaucourt biface quantitatively assessed as Novice will be qualitatively assessed as Expert, and vice versa. I do not expect such misclassifications to be prevalent due to the gulf in experience and sensorimotor control between the Novice and Expert knappers.

4.2.3 A Neurophysiological perspective

Understanding skill from an engagement perspective has been aided by neurophysiological research into the sensory realities of human involvement with the material world (Bruineberg and Rietveld, 2014; Friston, 2011, 2010, 2009; Friston et al., 2011, 2009, 2010; Gershman and Daw, 2012; Hawkins and Blakeslee, 2005; Malafouris, 2020a; Rietveld and Kiverstein, 2014).

Friston (2011, 2010, 2009) and Friston et al.'s (2011, 2010, 2009) 'free-energy' principle argues against traditional sensorimotor explanations of action and behaviour. Traditional approaches rely on the concept of utility and the notions of desires, rewards, and goals to explain behaviour (Gershman & Daw 2012). Moreover, traditionalists regard an action's utility and the likelihood of it occurring as separate from each other. They consider the

agent able to determine both independently of each other before making decisions (Gershman & Daw 2012). However, Friston and colleagues instead explain sensory behaviour as the act of reducing prediction error, or 'free-energy', through action, to minimise surprising encounters with the world (Friston 2009). They do not mention utility (Gershman & Daw 2012).

Proponents of 'free-energy' do not view actors as reflective assessors who make judgements regarding which action to take. Instead, they deem them to be unreflective predictors whose engagement with the world enables them to expect specific sensory outcomes of action in high-fidelity before they take place. Consequently, this influences which actions they perform. Sensory engagement can include extrasomatic tools in an individual's 'body schema' too (Bruner and Iriki, 2016; Clark, 2007; Hirose, 2002; Holmes and Spence, 2004; Iriki, 2006; Malafouris, 2020a; Maravita and Iriki, 2004; Miller et al., 2014). A notable example of this is the blind man's stick analogy, where perception of the world extends beyond the body (Hirose, 2002; Malafouris, 2020b, 2013, 2008).

The 'free-energy' principle proposes that through acting with, and in, the environment and registering the sensory outcomes of their actions, agents are then able to predict the sensory consequences of their actions in the future (Friston, 2011, 2010, 2009; Friston et al., 2011, 2009, 2010). This involves expecting the sensory inputs one perceives before the action takes place. As Hawkins and Blakeslee (2005: 158) put it: 'one's predictions not only precede sensation, they determine sensation'. To clarify, agents do not perceive each individual sense delivered in action but sample the sensory information that conforms to their expectations so that sensory surprise is avoided, and predictions are

fulfilled in a somewhat 'self-fulfilling prophecy' (Friston, 2009: 295). Consequently, one's perception is confined by their prior action which provides truthful predictions (of their own experience) that guide one's active sampling of their sensory input (Friston, 2009).

This description of the 'free-energy' principle is initially puzzling and necessitates further clarification. For example, how does an agent sample their sensory environment? How do they incorporate relevant predictions without some assessment of utility (Clark, 2013; Gershman and Daw, 2012)? Moreover, how does one develop their sensory predictive accuracy? If agents sample sensory information that conforms to their expectations, how do they reduce 'free-energy' (prediction error) by incorporating new and surprising senses into their active sampling of the world? Surely, conforming to sensory expectations precludes exploration and ignores sampling relevant information? Would this not lead to prediction error ('free-energy') not being minimised but remaining relatively constant and considerable (Friston et al., 2009)?

Friston and colleagues provide explanations for these questions by utilising the concept of *active inference* that works using a Bayes-optimal model (Friston, 2011; Friston et al., 2011, 2009). Active inference works by augmenting the range of an agent's expectations, regarding what sensory input to predict in a certain action, through experience in the world. This inevitably generates different permutations and consequences of action which can then be incorporated into an agent's repository of sensory experience. In other words, an agent's life experience of an action induces prior sensorial expectations, which guide perceptual sensory inference and therefore action in a loop (Friston et al 2011). Friston describes this as perception being 'enslaved' by action to provide truthful predictions that

guide active sampling of their sensory environment (Friston, 2009; Friston et al., 2011, 2010). The agent's action minimises proprioceptive, or kinaesthetic, (i.e., sense of self-movement and body position) prediction error by increasing the range of sensory inputs experienced, whereas an agent's perception reduces prediction error by altering predictions (Friston, 2009; Friston et al., 2010).

Predictions can only be altered with additional sensory input stemming from action. Therefore, when a sensory change is experienced that is significant enough to alter expectations, one starts to become more experienced as their range of expected sensory states of action is expanded and the number of unexpected states of action is reduced (Friston, 2011, 2010, 2009; Friston et al., 2011, 2009, 2010). Over time and with repeated action, the number of sensory surprises possibly encountered by the agent diminishes and, therefore, prediction error improves. Friston et al. (2010) explain that this process of active inference is an example of self-referenced learning (Maturana and Varela, 1972; Porr and Wöörgöötter, 2003) in which the 'actions of the learner influence its own learning without any evaluation process' (i.e., without conscious assessment of the utility of their action) (Porr and Wöörgöötter, 2003: 856).

This explanation is relevant to thinking about differences in skill-level in the Acheulean. A knapper with greater experience of an action can predict more accurately since they have encountered more permutations of performance and via active inference built a greater sensory repository for accurate prediction. Thus, they are more skillful. An agent with less experience is more susceptible to surprise and for more 'free energy' to be part of the action equation. This is because they do not have the same enlarged sensory archive of a more experienced agent. Thus, they are less skillful because they are less flexible.

To clarify, senses are not consciously 'selected' as relevant by the agent in reference to an internal representation, as traditional approaches may advocate. Instead, they 'emerge' as relevant based on embodied experience of prior action and an expectation of it predictably happening again. This is what Bruineberg et al. (2018) call 'action-readiness' to affordances and sensory conditions that allow them to attain a command of a certain situation. One's prediction of what is going to happen then necessarily sorts the senses that are attended to from the ones ignored (in action). If one is less experienced or has not encountered a particular sensory input, then that sensory information cannot emerge as 'relevant' because one does not have the memory of prior action and associated sensory consequences.

However, arguably an unskilled agent's action in the early part of their development is partly reflective since most of the sensory consequences of action have not yet been encountered. As a result, they remain hidden. As discussed in the previous section, one learns in a situated normative context and development of improved sensory prediction error must relate to achieving goals regarding a task that is shared by the group one belongs to. As Vaesan (2012) notes, complex tasks discourage individual learning but social learning helps to maintain motivation among novices and allows them to observe general actions for success, promoting individual learning. In other words, unskilled agents must trust in the utility of the actions they are performing and that they will ultimately develop the unreflective ability of more skilled individuals themselves. This view is different to the traditional approach in that it agrees with the 'free-energy' principle that utility assessments are not made in action. However, conversely it recognises that utility assessments are part of the wider skill development and socially

motivated learning scenario which help instil beliefs in beginners that certain actions will engender a greater likelihood for correct action.

Nevertheless, a learner's attention cannot be sustained on multiple simultaneous tasks because their prediction error is too high in one action to permit them the adequate freedom to attend to the sensory consequences of another one. Questionnaire participant 7 notes that the key mistakes beginners make correspond to an insufficient evaluation of the state of the biface and difficulty evaluating what is possible and what relevant actions should be attended to next (see Catalogue 2, response to Q.16b, pgs. 15-16). My own experience as a novice, as well as observation of other beginners, has shown that neophyte inability to overcome a problem, leads them to disregard other aspects of manufacture which they know they should focus on. Instead, they concentrate on improving their ability (or sensory prediction error) one task at a time. A good example of this is the common occurrence of beginners reducing the breadth and length dimensions of the biface in pursuit of thinning the biface to the level which they strive for. As questionnaire participant 3 notes (see Catalogue 2, response to Q.16b, pgs. 15-16), because they are continually unsuccessful in this pursuit, they produce small and thick handaxes with very sinuous edges that cannot be improved. I expect to find qualitative examples of this kind in the experimental and GZnivG assemblages which evidence the attempts of less skilled knappers to increase their prediction error in line with normative standards of action.

Additionally, I envisage that within my experimental assemblage Expert individuals will produce more consistent quantitative scores within a tighter range than Intermediate or Novice level knappers due to them possessing greater sensory predictive accuracy. Moreover, because of this, although I expect each experimental skill-level category to be

predominantly quantitatively distinguishable from each other, I anticipate that misclassification between skill-level categories will mostly occur between Novices and Intermediates. Experts in my experimental assemblage will only rarely be misclassified with lower skill-level categories due to their greater sensory experience in making bifaces that has led them to possess greater control over their engagement with the material. The insights gathered from my experimental analysis and the outcomes of my predictions will be used to assess GZnivG.

It is important to note here that the 'free-energy' principle does not contradict what was discussed in the previous section, that social influences impact intentions and constrain the emergence of actions. Proprioceptively, the agent still learns for themselves through repeated practice which develops their perception of relevant affordances in action; it is just that sociocultural constraints limit their engagement with the material world as they learn.

4.2.4 Plateaus, Dips, and Leaps (PDLs)

Most would attest that performance generally improves with practice. Though, Gray and Lindstedt (2017) and others (Gowlland, 2019; Ingold, 2000; Marchand, 2007; O'Connor, 2006) argue that this linear view of progression belies the truth of skill development. They use the concept of plateaus, dips, and leaps (PDLs) to illustrate the reality that skill progression does not directly correlate with the amount of practice. Their account of PDLs as a series of objective behavioural events mirrors well my own learning experience that will be presented as part of my experimental results in Chapter 7.

Plateaus are periods of stable suboptimal performance and can endure, even at a very high-performance level (Gray and Lindstedt, 2017). Dips occur when the learner's performance suffers as they explore their actions and new methods are invented and tested. Leaps, on the other hand, denote the implementation of a successful new method and an incremental jump above the path to a performance plateau with the previous method. Gray and Lindstedt (2017) explain that the methods developed during these dips and leaps are key to surpassing one's teachers and eventually acquiring 'extreme expertise'. However, leaps are not always guaranteed and failures in method invention can result in a reversion to old methods and continued plateaus.

Overall, although the PDL approach does not contradict the notion that the overall 'curve' of an individual's skill development increases with practice, it argues that it does not occur in a smooth incremental fashion correlating with additional cumulative practice. As Ericsson (2004) notes, there are long periods of stable suboptimal performance before improvements are discovered by the learner. The ups and downs of skill progression are covered in Wilson's (1999: 157) discussion with professional juggler, Serge Percelly, who said:

> 'You rehearse, you do a trick, you do another one, they're fine. You do it again, it's good, you do it again, it's better, and so on. [However] The next day you can't do it at all! You've practiced the whole day, you did it ten, twenty times the day before. The same thing keeps happening until it slowly starts to get better. The work gets less intense and the average gets slowly, slowly better'.

Percelly's testimony neatly demonstrates how there can be dramatic differences in performance from one day to the next, but the general trend can be one of improvement over time (Wilson, 1999). Therefore, it is essential to consider how dips or leaps may affect judgements regarding the skill-level of the knapper, especially when assessing handaxes in isolation from the associated waste from their production as I do in this thesis. In my experimental and archaeological investigations of skill-level, qualitative assessment is used as a recourse to resolve possible unfair quantitative classifications resulting from dips or leaps in performance.

The PDL account of becoming skilled also highlights the considerable determination and motivation it takes to overcome plateaus and dips in performance. Gray and Lindstedt (2017) explain that this drive is needed to overcome plateaus by league stepping methods – methods developed by an individual through practice and self-discovery of effective execution. However, Percelly admits that self-tuition constitutes the 'hard way' of becoming skilled and most who try will become apathetic and give up without proper support. This is consistent with Austin's view that learning to be more efficient requires guidance (Austin, 1976; Wilson, 1999). Austin (1976) explains that without tuition the learner will more likely be prone to repetition of the same unhelpful behaviours which will stunt development (Wilson, 1999).

Instead, more experienced help can aid intelligent rehearsal and increase the repertoire of practice when something goes awry. Indeed, as I outlined in previous chapters, numerous scholars have explained how handaxe manufacture is complex enough to require aided learning of some kind (Morgan et al., 2015; Muller et al., 2017; Shipton, 2019). Importantly, this external support does not give the learner knowledge because merely

communicating ideas to somebody is insufficient (Austin, 1976; Wilson, 1999). As Gray and Lindstedt (2017: 1847) note 'if knowing the goals and subgoals for performing a task was the only requisite for increasing expertise, then we would have a lot more extreme experts'. However, assistance from others facilitates new possible avenues for an individual to work it out for themselves, preventing recital of the same unhelpful tasks. Nevertheless, as questionnaire participant 6 notes, becoming skilled in the Acheulean would have taken time despite there being no shortage of good advice from others (see Catalogue 2, response to Q.18, pgs. 17-18).

My own experience knapping with stone-knappers, details the propensity for a learner's performance to leap when aided by expert instruction but severely dip when subsequently unassisted. The learner directly follows an expert's directives and, as such, executes how the expert wants them to. The level reached by the beginner constitutes a radical improvement in performance. However, it is not their own know-how they are using and, as such, the level reached cannot be sustained without direct instruction. If a novice is asked to perform immediately after being taught, there will be a significant dip which corresponds to a return to near the level previously held. From here, the learner knows what is possible but must discover for themselves how to attain the levels they attained with instruction. Effectively, the knapping decisions (e.g., where to prepare a platform, where to strike etc.) and bodily movements realised (e.g., angles of striking arm during thinning, position of handaxe during striking) under direct instruction must be found again and repeated by the learner to be able to reach the level they know is possible. My experimental investigation will test these observations of dips and leaps associated with tuition in more detail.

Gray and Lindstedt (2017) offer examples of the limitations of using other people's methods in the development of one's own skill too. They demonstrate that utilising another's approaches can initially expedite successful performance but only gets you so far before the strategies one tries to replicate become more specific to the other's knowledge, know-how, and specific route of skill development. As such, another's devices are less easily adopted and adapted and ultimately leads one down a dead end. To progress from this plateau, the individual then needs to invent their own techniques to progress further. Again, one needs to find their own way of doing things based on their personal sensorimotor experience, not just utilise the products of another's understanding.

The various permutations of skill development that highlight the non-linear, up-and-down nature of learning are noted by Gray and Lindstedt (2017). For example, a dip and subsequent leap in performance may signal the adoption, by the learner, of better goals and sounder methods for implementing those goals. Initial decline may relate to the beginner attempting variations of new approaches on variants of a particular task. Alternatively, a dip followed by a return to a plateau is more ambiguous and might signal fatigue or loss of motivation. Alternatively, it could denote a period of unsuccessful exploration that does not result in a new and more successful scheme. Furthermore, if there is no dip but a progressive leap, this may signify that a new routine was more efficient and straightforward to perform, with the cost of implementing it being not great enough to create a dip.

Overall, Gray and Lindstedt (2017) conclude that it becomes harder to leap once you get more skilled. This strengthens the conviction I have in my prediction that Experts, in comparison to Intermediates and Novices, will have a tighter range of performance.

Quantitative analyses of skill criteria will test this among my experimental and archaeological assemblages. In addition, although Gray and Lindstedt's (2017) research outlines that skill development among learners is not aggregative with each successive period of practice, I envisage that quantitative analysis of Novices within my experimental investigation will show improvements with repeated training. However, unlike Experts there will be regular dips in execution also and, as such, the range of performance of Novices will be more varied.

Moreover, I expect that advancements among experimental Novices will also occur in the event of tuition by another more experienced knapper. However, this progress will be ephemeral and consigned to direct tuition. As I explain above, this is because the Novice appropriates the experience and self-reflective knowledge of the Expert, yet tuition does not last long enough for the knowledge to be embodied in the Novice. Therefore, despite the Novice being more knowledgeable about what to do, they will ultimately suffer a prolonged post-tuition dip back to near their pre-tuition performance level. This is because they still do not possess the bodily know-how, predictive accuracy, flexibility, and grip over affordances of their more highly skilled tutor.

Testing whether tuition has a noticeably considerable effect on the performance of Novices in my experimental investigation is pertinent to understanding the significance of this variable in potentially misclassifying bifaces by skill-level. The outcomes of this analysis will serve to calibrate the quantitative and qualitative parameters of what is deemed possible by a beginner Understanding what Novices are capable of under instruction is useful when applied to archaeological material. One will be able to assess

the potential for tuition to complicate the confidence with which a skill-level classification can be given.

The above discussion of theoretical approaches to skill in general and in relation to handaxe-manufacture specifically, can be enhanced through detailing some ethnographic and experimental observations on the development of skill, the contexts of learning and practice, and the subsequent identification of variable skill-levels. Below I outline some of these.

4.3 Ethnographic case studies

It is unnecessary to detail here all relevant ethnographic studies of living craftspeople that relate in some way to understanding skilled action. Instead, I have selected a few examples to scrutinise the theoretical approaches presented above and provide additional insights into how we recognise the actions and products of variably skilled individuals.

4.3.1 Roux and Corbetta, 1989

Roux and Corbetta (1989) present two ethnoarchaeological studies of potters in India. The studies aim to provide reference material of craft specialisation and technical competence for use in the interpretation of archaeological data. Perceptual motor tests, gesture analysis and experimental data regarding wheel-thrown pottery in the first study are utilised to demonstrate its marked difficulty and the considerable effort apprentices must exert to master the technique. The second study attempts to establish the difficulties

of producing ceramic vessels of antiquity. The authors accomplish this by establishing a techno-morphological taxonomy based upon Indian potters' classifications, explanations and general testimony which is then assessed via further experimental production/investigation.

An important observation by Roux and Corbetta is that beginners were unable to simultaneously act in two motions. This perception is relevant to the notion of optimal grip (discussed above) – where skilled persons assess multiple affordances during action and possess the ability to concurrently attend to separate affordances (Bruineberg and Rietveld, 2014; Rietveld and Kiverstein 2014). As I mention above, my own experience as an inexperienced knapper and my observations of other beginners aligns with this observation.

Roux and Corbetta's study can be linked to Friston's 'free-energy' principle (discussed above) and a novice's poorer predictive abilities in action. Beginners are less able to predict the outcome of one action and how that result will affect subsequent actions. For example, this can quickly result in comminution, with the novice reducing the edges too far and making an irregular and poorly thinned handaxe which has had little attention paid to the production of a sharp cutting edge. I predict that my experimental investigation will indeed highlight that knappers with less experience work irregularly and inflexibly, and without a coherent plan. Overall, I expect that quantitative analysis will show Novice bifaces to be thicker, less symmetrical, and with more sinuous, or zig-zagged, edges.

4.3.2. Keller and Keller, 1996

Keller and Keller's (1996) ethnography of blacksmithing differs to Roux and Corbetta (1989) in that it aims to understand more about skill development through one of the authors' purposefully putting themselves through the rigorous process of apprenticeship.

Their account of apprenticeship is simultaneously from the viewpoint of observing the master craftsmen as well as from the perspective of the learner. The autobiographical aspect of the account touches upon facets of the learning process that cannot be accessed purely through watching. These relate to the emotional side of learning, the trials of failure, and the joys of success. Essentially, Keller and Keller recount the journey of knowing for yourself, highlighting the differences and variations in methods used to overcome problems by the novice in comparison to what they have observed.

Keller and Keller present a hybridised view on production and discuss the location of intentionality – in both space and time. Like Hutchence and Scott (2021), their account of skill progression helps to dissolve the dichotomy between internalist and MET views on production and intentionality outlined above. They do this by giving an accurate account of the various 'levels' at which intentionality resides, explaining that the specificity of one's intentions increases with engagement with the material.

At the topmost, generalised level, they outline that a skilled practitioner has an 'Umbrella Plan' which defines a goal for production and is made up of broad memories of action that need to be performed (without any specific content). These 'Umbrella Plans' are made up of 'Constellations' which are units of ideas, tools and materials that express

hypotheses for the integration of mental and material agencies for the achievement of a particular step in manufacture (Keller and Keller, 1996: 23). Akin to Friston's 'free-energy' principle then, Keller and Keller's notion of 'Constellations' are like predictions based on prior experiences that have been added to, transformed, and become more exact through repeated practice. They explain that the actions enabled via 'Constellations' simultaneously reproduce and redefine the potential for the 'Constellation' itself, the material artefact being produced, and the 'Umbrella Plan' from which it is derived (Keller and Keller, 1996: 23). In other words, a 'Constellation' is a hypothesis specifying the means for achieving an outcome within an overarching 'Umbrella Plan' (Keller and Keller, 1996: 106).

It is clear from this description that human intentions at both levels ('Umbrella Plan' and 'Constellation') are influenced by engagement with the material. In fact, Keller and Keller (1996: 118) note that 'the most thorough design and associate plan for action can never be sufficiently detailed, or adequately precise to anticipate everything that can happen during production. The process of production always has the potential to force changes in plans'. Thereby, they admit that craftspeople rely on the emergent characteristics of making to complete their plans, plans which are formed by their repeated engagement with the emergent characteristics of manufacture. Consequently, intentionality to produce a specific artefact does exist before manufacture, as internalists would view it, in the form of an 'Umbrella Plan'. It is made up of previous experience and projections regarding how one's emergent engagement with a material will unfold. Though, it is very generalised and lacking in precise content since the specific material engaged with in each case is responsible for supplying an equally critical aspect of the content of one's

intentions. Thereby, one can predict but not in full fidelity. As Keller and Keller (1996: 118) propose, ‘you have to follow the material’.

This can shed further light on the above discussion of Garcia-Medrano et al.’s (2019) statement that the ‘pathways to production’ and ‘end-goals’ are set in the mind of the knapper. Keller and Keller’s account shows that in one sense Garcia-Medrano et al. are correct, in that prior repetitive experience of engagement in the same task will amass a repository of expectations for future action and, at a very generalised level, will allow the knapper to envisage the ‘pathways to production’ and ‘end-goals’ they desire. Yet, future intentions only become content-specific during engagement with material. This is because, during making, the stone substrate will actively predicate the predictions of one’s actions. The specific content of emergent actions will then feed the repository of prior experience that furnish intentions at a more generalised level prior to active engagement. This explains why we see variations on a common goal at sites like GZnivG and Boxgrove.

In addition, Keller and Keller confirm, similar to Polanyi’s (1958: 4) acknowledgement that ‘we can know more than we can tell’, that expertise is non-declarative. They claim that it is ‘difficult for someone engaged in routine behaviour to articulate the tacit premises governing their performance’ (Keller and Keller, 1996 :111-112). Their explanation is that expertise is emergent over time and cannot be transmitted, and one must develop their own methods. This is indicated by Keller and Keller’s description of specific ‘Constellations’ and ‘Umbrella Plans’ which they admit do not provide any real proprioceptive or sensory understanding of where to place the body in relation to tasks. It is a person's own experience that facilitates making connections and grouping together

these constellations within ‘Umbrella Plans’. Keller and Keller outline that successful action is about making hypotheses, or predictions, which necessitates greater experience. An individual cannot predict without prior experience; accordingly, one is skillful if they can predict.

Therefore, as questionnaire participant 3 alluded to in response to the question of what makes a competent knapper (see Catalogue 2, response to Q.17, pgs. 16-17), I predict that if an Expert can envisage the outcomes of their actions more accurately and consistently, one should be able to discern quantitative and qualitative differences in comparison to those with less experience who cannot predict as exactly or reliably. I will test this hypothesis as part of my experimental investigation, before applying any insights to analysing differences between the bifaces at GZnivG.

4.3.3. Biryukova et al., 2015

Biryukova et al.’s (2015) recent ethnographic investigation of Indian stone-knapping looked at the kinematic solutions used by craftsmen of different skill-levels and demonstrated that expertise is marked by stability in task-crucial variables and flexibility in non-crucial variables (Williams-Hatala et al., 2020).

Specifically, they looked at arm joint movement variability regarding functional joint loadings (i.e., those that contribute to the position and angle of the hammer’s head) and nonfunctional joint loadings (i.e., those that do not support the position and angle of the hammer’s head). Biryukova et al. found that for every craftsperson, regardless of skill-

level, the functional component of joint loadings was less variable than nonfunctional ones. However, there was a marked difference among those with greater motor skill – they demonstrated greater stability in functional joint loadings, whilst their nonfunctional joint loadings were more mutable than those possessing inferior motor skill.

Given the above discussion of skill, the pattern that Biryukova et al. (2015) have identified among more skilled individuals makes sense. More stable behaviour among Experts suggests, in line with the 'free-energy' principle, that they can predict more accurately than Novices. Novices cannot as accurately envisage the sensory consequences nor the results of their actions and, as such, act more inconsistently. The fact that Experts are more highly flexible in non-functional actions also suggests that there is an embodied understanding of what is relevant to the task and what isn't, allowing for greater freedom in non-crucial variables. Again, in line with the above discussion on skill, this demonstrates that skilled individuals, due to their greater experience, have identified multiple ways to manipulate non-functional joint loadings that do not affect functional joint loadings in a way that disrupts the efficacy of the action.

The evidence presented by Biryukova et al. (2015) reinforces my expectation that, among my experimental assemblage, Experts will provide more consistent quantitative scores for skill-criteria related to the functional effectiveness of the handaxe such as edge straightness and refinement. This will be assessed as part of my experimental investigation.

4.4 Experimental archaeological case studies

The recent experimental kinematic study of modern knappers of variable skill-levels presented by Williams-Hatala et al. (2020) supports Biryukova et al.'s (2015) ethnographic observation that experts find various ways to exact the same outcome. Williams-Hatala et al. (2020) examined the manual pressure exerted on hammerstones of different weights when producing simple stone flakes using direct hard hammer percussion. They found that the cumulative pressures experienced by digits (fingers) when striking, and over time, was lower for experts. Specifically, they established that novices tended to hold hammerstones in a way that kept the pressure patterns across digits II-V (index to little finger) similar, whereas experts used digit II (index finger) more in the stabilisation of the hammerstone and used digits III-V (middle to little finger) more changeably.

The similar use of the thumb by all skill-levels demonstrates its importance in stabilising objects more generally, however, the exclusive emphasis on digit II by experts shows that they were able to act with energetic restraint compared to novices and intermediates. This suggests experts are more in tune with what is needed and what the consequences of their actions are likely to be. From a 'free-energy' perspective, this can be framed as experts being better able to accurately predict the energetic requirements of their body to achieve a particular outcome. The similarity with which novices use digits II-V suggests they are less certain of the outcomes of their action and cannot appropriately predict how much pressure or force to exert to achieve a specific goal. In contrast to the predictive confidence of experts, there is an suggestion of 'making sure' in the novices' behaviour, ensuring the core is secure and that they will hit it with the hammerstone.

The energetic forbearance exercised by experts, as well as their superior predictive abilities, is also demonstrated in the experimental studies of Nonaka et al. (2010) and Bril et al. (2010). Both studies reveal that novices produce more energy than is required to remove flakes, whereas experts adjust the force they use to detach flakes in line with other factors involved in the knapping process. Bril et al. (2010) showed that even with changes in the weights of hammerstones, the use of hammerstones not preferred by the knapper, and changing the knapping goal, experts were able to produce less kinetic energy whilst achieving a higher successful flake production rate. As Williams-Hatala et al. (2020) explain, this suggests experts were better able to approximate critical load to failure due to their greater experience with how much force is necessary for particular tasks (Dibble and Pelcin, 1995). This is further evinced in Nonaka et al.'s (2010) study where novice propensity to crush platforms was related to using too much force to remove flakes.

Nonaka et al. (2010) also show that only experts were able to accurately anticipate the shape of the product of flaking. Predictive accuracy is something questionnaire participant 3 deems a critical component of competence (see Catalogue 2, response to Q.17, pgs. 16-17). This ability, in conjunction with their capacity to alter their kinetic energy when detaching flakes from inconstant cores with differently weighted hammerstones, indicates that experts have a superior sensorimotor control of their actions (Wiilliams-Hatala et al., 2020). This translates into a greater predictive ability regarding how to act in relation to the affordances of the materials they hold in either hand (i.e., hammerstone and core) and to what they are trying to achieve (i.e., flakes) (Pargeter et al., 2020; Williams-Hatala et al., 2020).

Indeed, Pargeter et al. (2020) show, in relation to experimental Late Acheulean handaxe manufacture, that greater conscious predictive accuracy is improved among experts but not as significantly as Nonaka et al. (2010) suggest. They argue that it is perceptual motor control of the body that in fact allows the knapper to perform with greater striking accuracy and appropriate force resulting in increased flaking success rates (Pargeter et al., 2020; Williams-Hatala et al., 2020). Greater perceptual motor control importantly aids the management of speed-accuracy trade-offs in order to reliably detach large, invasive thinning flakes that are necessary to enable effective bifacial thinning and edging (Pargeter et al., 2020).

Pargeter et al. (2020) note that up to eighty hours of practice was insufficient to develop such sensorimotor control. Instead, novices adopted an alternative approach which targeted small flakes with acute platform angles to increase accuracy and flake detachment rates. However, they were unable to reliably detach the large thinning flakes indicative of greater skill. Pargeter et al.'s (2019) prior experimental study proposes that more than 200 hours of practice is needed to achieve such bodily control.

The adjustments made by experts in the pursuit of a manufacturing goal supports the discussion outlined above - those who possess greater grip (i.e., overall command) over their own engagement with material affordances in action can better assess the relevance of what needs to be done. Furthermore, they can be more flexible in moving their body (with tools) to produce the desired outcome. Again, this is not necessarily conscious and Pargeter et al. (2020) conclude that it is the development of basic sensorimotor flaking competence related to successful action execution that is necessary for successfully

producing Late Acheulean-style handaxes, rather than the ability to formulate viable flaking plans.

Pargeter et al.'s (2020) emphasis that conscious predictive ability may not be that relevant further plays into the 'free-energy' principles outlined above. Skill is emergent in action through the ability of the body's sensorimotor control to attend to the relevant stimuli and perform in an adequate way, not the conscious pursuit of predicting how one's preconceived intention will play out.

Both Pargeter et al. (2020, 2019) experimental studies utilise approaches that aim to quantitatively assess the skill-level of knappers making Acheulean handaxes. This allows them to detail the presence of PDLs (Gray and Lindstedt, 2017). For instance, an example of method invention among Pargeter et al.'s novice saw them target the removal of small flakes with acute platform angles. This method was conceived to improve their accuracy and get them closer to the standard of expert-made handaxes exhibiting more invasive flaking.

The results of Pargeter et al.'s (2020) study found that initial learning rapidly improved over the first 30 hours of training (3 assessments), after which a decrease in rate of improvement was seen towards the final assessment at 90 hours that effectively constituted a learning plateau. Throughout this plateau there is a persistence of occasional outliers exhibiting very low model scores, producing highly skewed distribution ranges, similar to the pattern previously observed in Eren et al.'s (2011) paper which looked at performance of an individual learner of 'intermediate' skill. Such insights play into the narrative that expertise is marked by stable yet flexible performance and reduced

variability in the objects produced. Therefore, I again expect to see a greater range in quantitative skill criteria scores among less skilled individuals than those with greater knapping experience in my experimental investigation.

By the end of Pargeter et al.'s study, it appears that participants were experiencing another period of skill improvement. The authors projected how long their learning curves, with the current trend, would take for each participant to reach a metrically perfect score of '5'. Some were close to achieving it at 90 hours, others would have taken upwards of 400 hours given their poor progress. This demonstrates the motivation necessary to improve towards an intermediate level whilst also alluding to disparities in innate ability. Such determination allows the learner to deal with the PDLs they go through over the course of their development. Moreover, it highlights how important external support is in the development of an individual's ability in relation to learning a complex skill, not only in terms of instruction but in facilitating and aiding the novice to continue to practice and learn for themselves.

4.5 The relevance of the 'social' in discussions of skill

So far, this discussion and overview has shown that previous approaches to skill (traditional internalist, emergent and neurophysiological) have been mainly focused on trying to explain skill from the point of view of the individual interactions between agents and their material. Skill is developed personally and knowing for yourself is important. However, there are external sociocultural influences on perception, action and the development of an agent's sensorimotor contingency which are relevant to discussions of

skill also (Bruineberg and Rietveld, 2014; Rietveld and Kiverstein, 2014). Moreover, the sociocultural human 'form of life' is instrumental in aiding the motivation of learners to continue with difficult tasks which they regularly fail at or plateau in (Gray and Lindstedt, 2017; Pargeter et al., 2020, 2019). Social effects on behaviour are potentially important then for investigating skill-level variability in the archaeological record - they can be used as a control variable for determining relative ability between individuals who adhere to similar constraints on action.

An archaeological example of social constraint on action in the Acheulean has been proposed for Boxgrove, which exhibits an assemblage of handaxes with a consistently applied technological characteristics and shape design across multiple generations (García-Medrano *et al.*, 2018; McNabb, 2017; Pope, 2002; Roberts and Parfitt, 1999). In the following chapter, I propose a similar phenomenon at Gouzeaucourt Level G.

The above discussion emphasises that a proper exploration of skill necessitates the acknowledgement of sociocultural influence on the interactions between individuals and their material environment. Failing to do so, even if trying to understand skill from an embodied perspective, plays into the narrative of the technological definition of skill outlined in the introduction to this chapter. By not actively engaging with the fact that skill-level is relative within a sociocultural context one inadvertently judges skill at a universal level. At this level, one compares individuals based on what they can do but assumes the goals were identical across the Acheulean. Skill should not be comprehended this way. It is always relative to what is trying to be achieved (i.e., the sociomaterial constraints connected to the group an individual belongs to) regarding variations in handaxe manufacture.

Ascertaining what constitutes the sociomaterial constraints in a group is pivotal in being able to determine what is skilful on behalf of the knapper. Competent participation in a socially shared skillful act requires both a communally negotiated idea of what constitutes acceptable performance, as well as its reproduction by individuals in practice (Rietveld and Kiverstein, 2014; Wenger, 1998). Acceptability can relate to both the final product and the way in which it was achieved. In turn, understanding what is competent from a functional and social perspective generates a blueprint for deciding how skillfully an individual has adhered to sociomaterial constraints (Rietveld and Kiverstein, 2014; Wenger, 1998). Within sociomaterial limitations, there exists a certain amount of freedom to make mistakes among learners. Therefore, there should be a visible skill spectrum related to any task that incorporates a shared repertoire of actions and ideal outcomes. Therefore, GZnivG has been chosen to investigate differences in skill-level among knappers who are closely associated regarding the sociomaterial constraints on handaxe manufacture.

4.6 The skill spectrum: the theoretical reality

In this chapter, I have summarised some of the ways in which the topic of skill has been thought through in relation to crafts and everyday actions. I have discussed what constitutes proficient action and how expertise is acquired. I have also outlined how these approaches to understanding skill have influenced my methodology and hypotheses, presented in chapter 6, and how they will be used to investigate skill-level differences in the Acheulean.

It is clear from the above discussion that one should be theoretically able to discern the products that were made at the beginning of skill development from those made towards the end. If we extend this to a group of practitioners of the same craft who are all at varying points along a path to becoming skilled, we should be able to recognise the products made by beginners, those possessing moderate experience, and by veterans of the craft. This is the theoretical reality that I wish to investigate at GZnivG.

In the following chapter (5), I outline the rationale for conducting experimental investigation for interpreting skill-level differences at Gouzeaucourt Level G, as well as provide an explanation as to why GZnivG is suitable for the purposes of my thesis. After this, in Chapter 6, I present my Methodology which outlines the analyses I conducted on both experimental and archaeological biface assemblages alongside what I hypothesised would be the outcomes of them. I will then conclude the thesis by presenting my results and my interpretation and discussion of them in relation to my hypotheses.

5. Chapter 5 – Research Rationale and Method of Approach

5.1 Introduction

This chapter focuses on providing a rationale for the experimental and archaeological investigations that constituted my examination of skill-level differences between bifaces. My experimental rationale and method of approach is presented first. It centres around demonstrating methods for discriminating between modern knappers of variable skill-levels. Following this, my rationale for using GZnivG to investigate skill-level differences in the Acheulean is given. My reasoning hinges on the morphological and technological consistency of the knappers at GZnivG. My method of approach includes assessment of whether, as previous studies attested, handaxe shapes at Gouzeaucourt were consistently similar (Field, 2005; McPherron, 1994; Soriano, 2000). In addition, I examine whether this regularity was a result of a shared morphological objective, or a consequence of resharpening at Gouzeaucourt. To do this, it was necessary to perform shape analysis which accounted for size differences and removed any allometric biases (Crompton and Gowlett, 1993; Lycett et al., 2006).

As I mentioned in the previous chapter, I reasoned that if one can demonstrate a shared morphological objective (i.e., reduced morphological variation and common manufacturing techniques) among an assemblage of closely associated handaxes, it can operate as a control variable with which to investigate skill-level differences between those handaxes.

5.2 Experimental Investigation

The aim of my experimental investigation has been to generate quantitative data and qualitative observations regarding known skill-levels among modern flintknappers which could then be applied to the analysis of the archaeological handaxe assemblage from Gouzeaucourt. The motive for choosing to experimentally investigate skill is that there is an extensive history of researchers and knappers who have studied the shared and specific behaviours involved in knapping a variety of stone-tool traditions. They have highlighted observable quantitative and qualitative indicators of skilful action (Bril et al., 2010; Darmark, 2010; Eren et al., 2011; Ferguson, 2003; Hecht et al., 2015; Herzlinger et al., 2017; Khreisheh, 2013; Khreisheh et al., 2013; Newcomer, 1971; Pargeter et al., 2020, 2019; Shipton and Nielsen, 2018; Torres and Preysler, 2020; Winton, 2005).

One study particularly relevant to my own is Pargeter et al.'s (2019) analysis of modern novice knappers learning handaxe manufacture. The authors document how skill-level improves over time and present experimental findings that demonstrate the proficiency of an individual's performance of certain actions can be evinced quantitatively through studying the products they make. They did this by comparing performance during manufacture with what was ultimately produced as individuals gained experience. This research has been influential in designing my experimental investigation of archaeological handaxes; however, my method of approach goes further by using experimental data directly in the interpretation of skill-levels of Acheulean biface makers. As far as I am aware, this is something that has not previously been attempted.

Using similar methods of analysis to Pargeter et al. (2019), this thesis focuses on what is solely demonstrated by the handaxes themselves and not the waste from their production. The rationale for this stems from the fact that the end product is a manifestation of an entire suite of actions indicative of the relative ability of the knapper. Although associated waste-production material can yield additional insight into the aptitudes of knappers in ideal circumstances (i.e., through refitting sequences and/or *in situ* artefact deposits), these are extremely rare in the archaeological record (e.g., Boxgrove GTP17 Horse Butchery Site (Pope et al., 2020b)). Therefore, concentrating on handaxes here reflected better the material reality of most of the Acheulean archaeological record and what researchers investigating skilled practice must work with.

Much like Gowlett's (2006) GLOBFELTS, which catalogued some of the key similarities exhibited by handaxes across the spatiotemporal extent of the Acheulean, experimental replication, and analysis of archaeological handaxes has revealed some of the general behaviours necessary for skilled production of handaxes. These relate to what the functional requirements of the tool were, as well as which behaviours were more complicated to perform and required more practice. By comprehending which handaxe features are symptomatic of greater or lesser skill, variable abilities of both contemporary and past knappers can be recognised. However, quantitative analysis alone can sometimes obscure the interpretation of skill-level in making handaxes. For example, raw material effects can adversely impact the results achieved by the same individual. In addition, the preferred style and shape (i.e., pointed, or ovate) inevitably affects what was attempted and what was achievable in terms of observable skill-related metrics. Consequently, qualitative indicators of skill, such as instances of battering, poor edge control, or repeated step/hinge fractures, were included to avoid an over reliance on quantitative data.

Therefore, although the objectives of handaxe manufacturers must have been different throughout the Acheulean (Soriano, 2000), experimental investigation can deliver generalised insights into the common differences between skill-levels. It is a necessary foundation to build from when interpreting differences between archaeological handaxes. This is because the context of manufacture for experimentally made biface specimens is known and allows researchers to understand the general rules governing observable skill-level differences, as well as the caveats that exist. As will be explained in the following section, this generalised method of approach for determining skill-level differences among contemporary knappers is most appropriately applied to the archaeological record when used to compare bifaces within specific handaxe assemblages where the goals of manufacture were consistently replicated, and not between assemblages separated in space, time, and intentions.

5.3 Archaeological Investigation

Like Leroyer's (2016) assessment of skill-level of bifaces from Boxgrove, identifying the products of specific individuals at GZnivG was not attempted. Instead, the purpose of this study was to use insights from experimental research to identify the products of categories of individuals (i.e., Novice, Intermediate, and Expert) and to assess their relative contribution to the assemblage at Gouzeaucourt Level G.

As previously mentioned, investigating differences in the aptitudes of knappers in the archaeological record gains greater significance when comparing handaxes that are

associated within an assemblage and made by hominins who share ideas regarding biface making. Skill is, and would have been, understood relationally between members of an associated group or tradition. It is thus difficult to compare between assemblages, especially those assemblages significantly spatio-temporally removed from each other. Indeed, inter-site comparisons instead treat skill as a measure of technological complexity that can be compared between handaxe traditions and generalised for all members of each tradition (i.e., Later Acheulean handaxe makers who utilised more advanced knapping techniques, such as platform preparation, were more skillful than hominins in the Early Acheulean who made handaxes without preparing platforms). Therefore, focusing on interrelated groups of handaxe-makers who adhered to similar, or mutual, behavioural repertoires like at Gouzeaucourt Level G, will allow data to be obtained regarding the skill-level categories present within an Acheuelan group. Determining the proportion of each category has the added potential to yield important insights regarding site function too.

Consequently, approaching the topic of skill at the level of individual assemblages is a vital first step. Individual groups, even within the same broader tradition, hold their own shared history of learning and manufacturing intentions. It is within these groups, and not between them, where gradations of skill-level based on level of personal experience can be identified and used to generate insights into the makeup of those contributing to an assemblage. The relative skill-levels within these groups are contextual to the individuals and should be separated from the technological tradition they belong to.

Yet, demonstrating that handaxes belong to one assemblage from the same clearly defined stratigraphic horizon is often hindered by multiple factors relating to the antiquity of the

Acheulean record. Post-depositional and taphonomic processes (i.e., fluvial, slope, or anthropogenic movement of handaxes out of and into assemblages) often affect the integrity of an assemblage and can lead to the complete obliteration of context and association between artefacts. An extreme example of this is the formation of palimpsest assemblages of handaxes. In Britain, they are usually attributed to glacio-fluvial action (Hosfield, 2005; White, 1998) and preserve no original context or details of initial deposition of artefacts. Most handaxe assemblages, although certainly not palimpsests, are affected by some degree of natural and anthropogenic post-depositional and taphonomic disturbances. It is rare to find high-resolution spatial integrity and *in situ* preservation. Hence, it is uncommon to find a behaviourally direct, primary association between handaxes that constitutes a snapshot in time (Roe, 1980). Theoretically, such a snapshot would allow greater confidence in knowing the stone tools being analysed were deposited by members of the same group. Instead, there may be assemblages made up of multiple occupations by groups over time. This does not preclude the possibility that numerous respective contributions to an assemblage were made by knappers that shared a collective goal, as Leroyer (2016) showed at Boxgrove. Nonetheless, the joint ambition needs to be demonstrated and cannot be assumed. These considerations must be made in order to evaluate the association of the bifaces in an assemblage.

For the purposes of investigating skill differences among knappers that adhered to a communal concept of production, an assemblage of one (or more likely multiple) occupation(s) (at the same location) where taphonomic effects were small and biface morphology was homogenous was required. This was to demonstrate the genuine association of the artefacts. Morphological and techno-functional similarity of the handaxes, as well as a unified technological strategy of flake manufacture, from

Gouzeaucourt Level G were the principal reasons for choosing this unit to test ideas on skill in the Acheulean (Field, 2002; McPherron, 1999; Soriano, 2000). Level G also has the benefit of not being greatly affected by post-depositional processes (Lamotte, 1992; Soriano, 2000; Tuffreau and Bouchet, 1985). Consequently, before my methodology for investigating skill differences among the knappers of Gouzeaucourt Level G is presented, it is necessary to first provide a background to this handaxe assemblage.

5.4. Gouzeaucourt Level G (GznivG)

5.4.1. Contextualising Gouzeaucourt

Gouzeaucourt Level G has been indirectly dated to MIS 8 (300 – 243 kyr BP), a period of considerable technological development and climatic variability. Technologically, the Early Middle Palaeolithic of north-western Europe tracks significant change over the course of MIS 8, as well as increased assemblage variability in the artefacts produced (Malinsky-Buller, 2016). At the onset of MIS 8, bifacial production continued from MIS 9 as the main mode of lithic tool production, though instances of flake tools and prepared core technology began to become more widespread (Hérisson et al., 2016; Malinsky-Buller, 2016; Scott and Ashton, 2010; White et al., 2006). This pattern changed by the end of MIS 8 and the onset of the MIS 7 interglacial, with most lithic assemblages being dominated by Levallois reduction (Malinsky-Buller, 2016). However, handaxe manufacture still occurred during MIS 7 and 6, just considerably less frequently. MIS 8 is interpreted as a complex cold stage, also characterised by considerably fewer instances of hominin occupation in more northerly regions (Hérisson et al., 2016).

This archaeological pattern is assumed to reflect fluctuations in human occupation of north-western Europe due to climatic change over the period from MIS 9 to MIS 7 (337-191 kyr BP) (Hérisson et al., 2016; Malinsky-Buller, 2016; Scott and Ashton, 2010). MIS 8 was a cold stage that likely generated a mammoth-steppe environment which engendered increasingly patchy dietary resource distribution (Guthrie, 1990; Malinsky-Buller, 2016). Increased spatial separation of resources was previously argued to have promoted isolation of hominin groups during MIS 8 and greater variation in technological repertoires (Hopkinson, 2007; Hopkinson et al., 2013; Malinsky-Buller, 2016). However, Bocquet-Appel and Tuffreau (2009) suggested that maintenance of a continued high carrying capacity during cold periods like MIS 8 was the leading driver of technological diversity. The technological change from predominantly handaxe-dominated assemblages during the main span of MIS 8 to Levallois-dominated at the terminal phase of MIS 8 is widely thought to be spurred by climate oscillation (Hérisson et al., 2016; Malinsky-Buller, 2016; Scott and Ashton, 2010). This 'technological break' is believed by some to reflect a population, or cultural, discontinuity driven by a hiatus in occupation of northern Europe during intensely cold pleniglacial (ice-dominated) conditions which was then followed by subsequent recolonisation during the amelioration of the climate towards the end of MIS 8 (Malinsky-Buller, 2016; Scott and Ashton, 2010).

Whether human occupation of northern Europe throughout MIS 8 reflected the sporadic presence of hominins during warmer spells in what was generally a very cold place to live, or a more continual occupation, is still much debated (Bates et al., 2014; Scott and Ashton, 2010; White et al., 2006). However, archaeological evidence of the variability in stone-tool equipment certainly suggests the possibility that hominin groups and ideas

were moving around, and their technological responses and solutions altered substantially over the course of MIS 8.

What is undisputed, is that the levels at Gouzeaucourt were more representative of earlier MIS 8 assemblages in that there was a high propensity of biface production. In addition, flake production was Clactonian-like, and not Levallois (Boëda, 1995; Soriano, 2000). Soriano (2000) notes that GZ bifaces were more indicative of Acheulean than Mousterian forms. More specifically, Tuffreau et al. (2008) observed that the predominantly ovate GZ bifaces illustrated the transition from MIS 9 to 8. During this period, there was an observable change in biface morphology from predominantly pointed as well as mixed assemblages (pointed and ovate forms) prior to MIS 8 to more ovate-dominated assemblages in MIS 8. This pattern contrasted to Britain, where MIS 8 assemblages continued to be predominated by pointed handaxes at sites such as Cuxton and Furze Platt (Bridgland and White, 2015). This could suggest the introduction, or development, in MIS 8 of new biface-making traditions into northern France. This is especially the case when the increased transported and/or curated nature of the French material in MIS 8 is considered versus the largely non-curated signature of pre-MIS 8 French sites (Bridgland and White, 2015; Tuffreau et al., 2008).

Yet, in contrast to British sites of this period, Gouzeaucourt sustained a consistent approach to biface manufacture over multiple levels. Moreover, this was also visible at the nearby site of Longavesnes that dated to the same time and displayed very similar handaxe shapes (Ameloot-Van Der Heijden, 1993; Tuffreau et al., 2008). Therefore, within the wider picture of change and diversity, there were periods of technological stability that arguably reflected shared, socially agreed upon solutions to raw material

availability (e.g., curated biface manufacture of predominantly ovate forms at Gouzeaucourt over time).

5.4.2. Gouzeaucourt overview

Gouzeaucourt is an open-air site excavated as part of rescue excavations from 1985 to 1993, located within a complex of ancient loesses at the bottom of the eastern slope of the Vallée du Muid in the upper basin of the river Scheldt, northern France (see Figure 1.1 for location) (Lamotte, 2001, 1992; Scott and Ashton, 2010; Tuffreau et al., 2008; Tuffreau and Bouchet, 1985). The archaeological sequence was located within a karst depression, typical of the eastern loessic region of northern France, and exhibited multiple levels of occupation, each containing numerous handaxes and flake tools (Sommé, 1977; Tuffreau et al., 2008). The raw material sourced was predominantly in the form of small nodules of irregular morphology, often altered by frost cracks, and lay in a secondary position in nearby slopes, derived from chalk (Soriano, 2000). Bifaces were produced by both reduction of an entire, or fragmented, nodule, or on flakes (Soriano, 2000).

The handaxes, and flake tooling, from each layer exhibit striking technological and typological unity (Soriano, 2000). Bifaces are predominantly ovate with a high propensity of tranchet sharpening and display reduction via resharpening (Field, 2005; Lamotte, 1992; McPherron, 1999, 1994; Soriano, 2000; Tuffreau et al., 2008; Tuffreau and Bouchet, 1985). Interestingly, some levels exhibited *in situ* handaxe manufacture (Levels H and R) whilst others revealed the importation of handaxes produced offsite (Level G) (Scott & Ashton, 2010; McPherron 1994). Level (Niveau) G is my principal archaeological layer of analysis and yielded over 150 bifaces. Pedogenesis in the dark

brown clay loam that sat above it has been attributed to an MIS 7 palaeosol, therefore Level G can be indirectly dated to MIS 8, c. 300 kyr BP (Lamotte, 1992; Tuffreau et al., 2008).

The Gouzeaucourt level G (and H) bifaces have likely only been subjected to minor disturbance as they were deposited on a flat surface (Soriano, 2000; Tuffreau & Bouchet, 1985). Handaxes represented just 2.1% of lithic remains and 10% of the tools recovered from GZnivG (Lamotte, 1992). We can be certain that the bifaces from this level were produced off site from reduction and re-sharpening analysis, as well as the lack of evidence of refitting from the considerable number of lithic debitage and flake tools found in association with the bifaces (McPherron, 1994; Scott and Ashton, 2010; Tuffreau et al., 2008). Soriano (2000) also challenges the uniqueness of the deposit at GZnivG due to the identification of some handaxes exhibiting a double patina. This, he explains, means Level G most probably accrued from several occupations separated in time (Soriano, 2000). In addition, Soriano (2000) noted that the flint nodules worked *in situ* in Level G were of unsuitable morphology and quality for the shaping of bifaces found there. This, he rationalised, may explain the delocalisation of biface production. However, the distance over which the handaxes were moved remains unknown.

It is clear from the above that post-depositional action at Level G did not contribute to removing evidence of by-products associated with biface manufacture. Although such evidence could have helped with the analysis of skill-levels at the site, it was never part of the deposit and was located where the handaxes were transported from. Therefore, at the many sites like this, the only possible way to investigate skill-level is through analysing the bifaces themselves. Yet, some bifaces do show secondary signs of

modification with flake removals more consistent with later prepared-core technologies (Lamotte, 1992; Soriano, 2000; Tuffreau and Bouchet, 1985). This is a potential area of concern regarding the removal of observable evidence related to skill-level. One explanation for this pattern is that once the bifaces were deposited by their original makers, they were re-used by the same or different groups in an altered way – occasionally removing flakes from handaxes and transporting them offsite. It is possible that GZnivG acted as a pre-prepared synthetic outcrop that later populations with new technological traditions could exploit – i.e., removing flakes for other tasks. Such an explanation is less surprising than perhaps at first glance. Handaxes, at this time, were already being curated and regularly moved across the landscape (Tuffreau et al., 2008), presumably with the intent of conserving raw material in an increasingly harsh climatic environment (Hérisson et al., 2016; Malinsky-Buller, 2016; Scott and Ashton, 2010). Thus, the later change in use of handaxes as cores is consistent with this more economical use of raw materials in France from MIS 8 into MIS 7, as outlined above.

This later reworking visible on some GZnivG bifaces occasionally culminated in the removal of substantial portions of the original biface. Where the secondary reworking of the handaxes was extensive, I have omitted these bifaces from my analysis due to the reduced information that can be gleaned regarding manufacturing skill. This is because, some of the later flake removals have removed the ability to measure several quantifiable skill criteria, presented in Chapter 6, of the original handaxe. A relevant example of the impact of this later reworking is demonstrated by biface number 8015(9) (Figure 5.1). In total, 20 bifaces from GZnivG have been omitted from my analysis of skill-level at the site for this reason. However, they are included in my shape analysis below.



Figure 5. 1. Example of secondary reworking on Biface 8015(9). Significant removals affect biface shape and obliterate numerous areas for assessing skill-level.

Despite some later reworking, the bifaces, and flake tooling, recovered from across the several archaeological levels at Gouzeaucourt were strikingly technologically and typologically similar, whether made on site or brought in from outside (Lamotte 1992; Mcpherron 1994, 1999; Soriano 2000; Tuffreau & Bouchet 1985). In section 5.4, the outcomes of a comparative morphological analysis of handaxes from Level G to handaxe assemblages from other Gouzeaucourt levels (H and R) and from two other Acheulean sites (Boxgrove and Broom Pits) are presented. I demonstrate a marked morphological homogeneity within Level G, as well as across the site. This is notable and suggests, as others have indicated (Field, 2002; Soriano, 2000; Tuffreau and Bouchet, 1985), that the handaxes found across the levels, including level G, were variations on a shared, similar bifacial template of manufacture held by hominins who repeatedly occupied the area.

I agree with Soriano (2000) that this morphological pattern was due to socio-cultural influences on manufacture and not, as McPherron (1999, 1994) argued, solely the result of utilitarian approach to raw material conservation through re-sharpening. This claim, as mentioned above, is strengthened by very similarly shaped and aged handaxes recovered from the nearby site of Longavesnes (Ameloot-Van Der Heijden, 1993; Tuffreau et al., 2008; McPherron, 1994). Moreover, when allometric effects on shape were accounted for via size-adjusting and standardisation, which is dealt with in greater detail in section 5.5 below, the handaxe shapes from GZnivG are very similar regardless of size (see Figure 5.2). Therefore, in the next section, I contest the utilitarian interpretation of the morphological pattern at GZnivG in more detail.

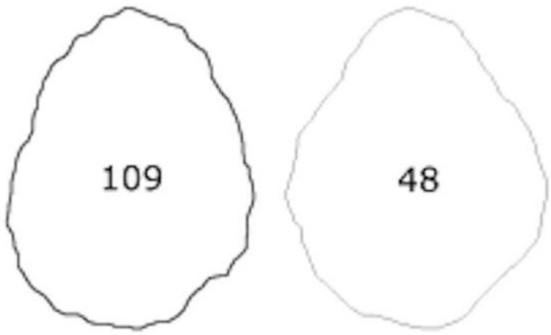


Figure 5. 2. Consistency of biface shape at GZnivG, regardless of size. Two bifaces from GZnivG, size-adjusted and standardised. Numbers reflect real length in mm.

5.4.3. Challenging a utilitarian perspective of Gouzeaucourt

In Chapter 3, McPherron's reduction hypothesis was introduced as a utilitarian approach to understanding biface variation which challenged the traditional model that handaxes

were intentionally shaped. Here, the limitations of this approach are discussed with reference to Gouzeaucourt.

McPherron's model posited that hominins reduced bifaces from originally larger pointed forms down to smaller ovate forms towards the end of their use-life. This theory was based on several assumptions. One, once a biface was made, it would be reused and reduced. Two, reduction would be focused at the biface tip, because the tip was worked before the base (as evidenced by handaxes with cortical bases). Three, bifaces started off larger, pointed and less refined and became more rounded and refined as their tip length was reduced (McPherron, 1999). Beyond a certain resharpening threshold, however, bifaces could not be thinned any further which led to a secondary associated increase in thickness relative to length and breadth (i.e., refinement) as bifaces became smaller.

McPherron applied this logic unanimously and, as such, suggested that Gouzeaucourt's ovate bifaces represented the middle and late stages of reduction via resharpening. He used this functional perspective to argue that similarity in biface shape was an unintentional result of re-sharpening at this site and elsewhere. I, however, contend that the contrast between the variable intensity of reduction, the homogeneity of the forms produced, and techniques used at Gouzeaucourt may also be indicative of the existence of a socially significant morphology.

Broadly speaking, I dispute the notion that there is a singular universal trajectory for biface manufacture and use. Not all bifaces were reused and resharpened after manufacture. Indeed, numerous scholars (García-Medrano et al., 2019; Leroyer, 2016; Shipton and Clarkson, 2015; Shipton and White, 2020) have demonstrated that handaxes

at Boxgrove were not overly curated or modified after initial manufacture. I would agree with Shipton and White (2020), Garcia-Medrano et al. (2019) and McNabb (2017) that morphological homogeneity at Boxgrove, and the fact they were not resharpened, suggests that the initial form was intentional. Therefore, without convincing evidence for reduction from point to ovate at a site, it cannot be assumed to have been the case. It could equally be the case that initial forms were ovate and were subsequently resharpened. Moreover, re-sharpening and deliberate planform shaping are not necessarily mutually exclusive, as Figure 5.2 helps to illustrate. Furthermore, as Figure 5.3 shows, there was no significant relationship (*rs* (161) 0.1, p=0.895) between shape (as determined by Roe (1968)) and length in the Gouzeaucourt Level G assemblage as McPherron's reduction hypothesis would suggest. Oddly, some of the most pointed bifaces were the shortest in the assemblage from Level G.

Additionally, McPherron's (1999, 1994) comparison of Cagny-La-Garenne (CLG) and Gouzeaucourt (GZ) highlights how the assumptions of his paradigm were unsatisfactory. CLG exhibited a wider range of both larger and thicker pointed forms and smaller and more refined ovate forms, whereas GZ revealed 'narrowly defined' forms, technologies and sizes constituting around 90% ovates (McPherron, 1999). As per his hypothesis, McPherron interpreted the similarity between 'more reduced' forms at CLG and 'less reduced' forms at GZ as indicative of the singular reduction trajectory previously mentioned. Crucially though, tip length was the only evidence McPherron used to measure reduction intensity. This would be less of an issue if CLG represented an *in situ* handaxe assemblage, however, it comprised a mixed accumulation recovered from fluvial deposits. Its mixed nature raises doubts as to the efficacy of the conclusions McPherron proposed regarding the reduction trajectory at both CLG and GZ.

Another key consideration is that CLG has been relatively dated to Middle Terrace II of the Somme valley terraces, around MIS 12/11 (478-374 kyr BP) and is thus in parts at least 74,000 years older than GZ (MIS 8) (300-243 kyr BP) (McPherron, 1994). Moreover, CLG is of course also affected by significant post-depositional processes whereas GZ is not. The difference in age of the two sites is relevant given the differences in hominin demographics, climate, and technology between the two periods. The palimpsest nature of the CLG assemblage and its high morphological variability were both factors overlooked by McPherron. He regarded reduction by resharpening as a defining criterion. However, as I have noted, handaxes can be intentionally made ovate from the start. Therefore, CLG may instead consist of separate expressions of handaxe traditions broadly dated to MIS 12/11 – some that were larger and pointed and others that were small and ovate. In line with this interpretation, Soriano (2000) noted that consistently applied techno-functional aspects of production, as well as the marked regularity in biface shapes, among most of the sample he studied from level G, indicated similar intentions were shared by associated knappers.

In light of the above considerations, arguing that the handaxe assemblages from Gouzeaucourt, which are highly homogenous and not overly affected by post-depositional action, is the result of reduction from more pointed forms becomes more difficult. Therefore, although resharpening resulted in more rounded and thicker (relative to their length and breadth) handaxes as they became smaller (McPherron, 1999, 1994), the consistent handaxe shape and techno-functional flaking patterns (Soriano, 2000) at Gouzeaucourt can plausibly be argued to relate to a purposeful starting (ovate) shape in mind.

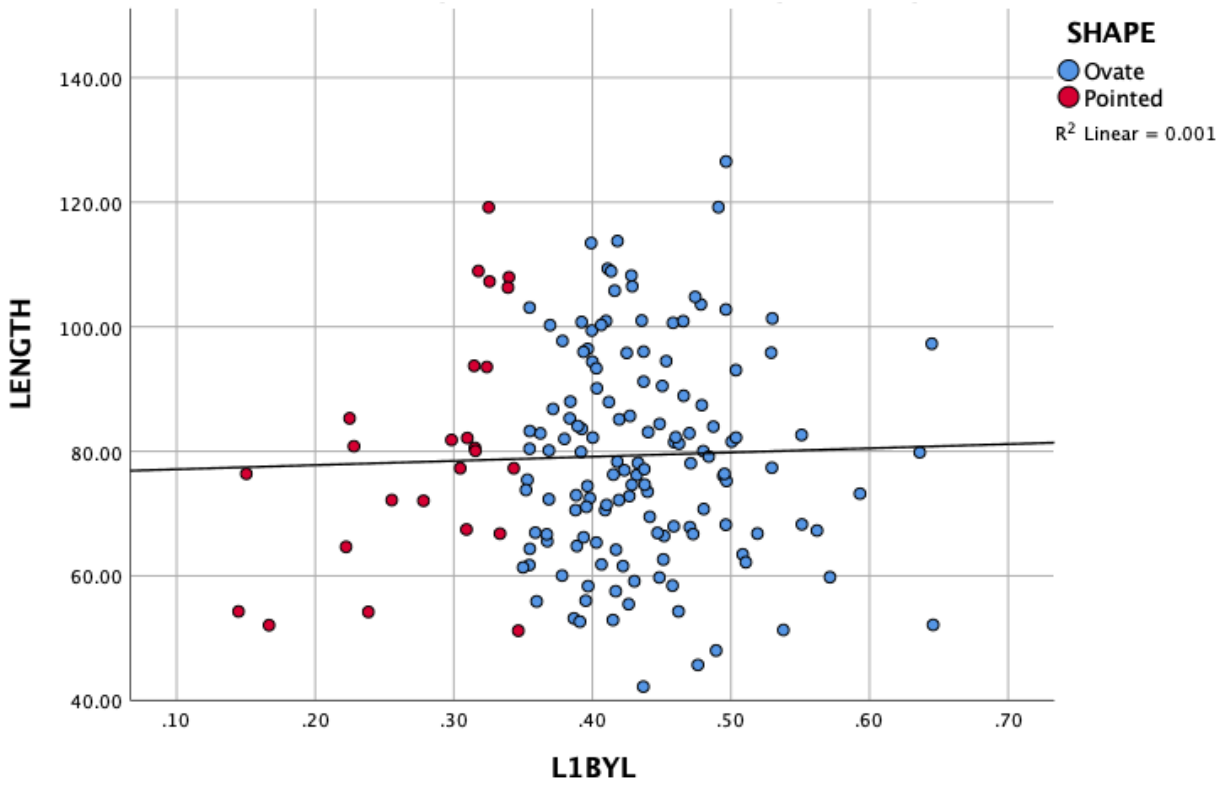


Figure 5. 3. Grouped Scatter graph correlation of Length vs L1/L at GZnivG sorted by biface shape. (L1/L is Roe's shape measurement distinguishing Ovate from Pointed bifaces).

Indeed, McPherron (1995) explained that ovates became far more refined than the points they supposedly started off as, mainly due to a change in reduction strategy once the tip length decreased to a more ovate shape. He rationalised that as this happened the reduction strategy, although still focused on the tip, would begin to incorporate reduction from the entire perimeter of the working edge. However, there is no evidence at Gouzeaucourt for a change of strategy. All the handaxes from each of the Levels at the site were worked around the entire operational edge. Furthermore, as McPherron admitted, the original size of the nodules was, as Soriano (2000) pointed out, likely to be small and irregular, or made on flakes. Therefore, at Gouzeaucourt it was more probable that the original form of the handaxes in the numerous levels was ovate.

At GZnivG, the behavioural repertoire was one of reduction of handaxes in a regular shape, despite variable reduction levels and size differences (see Figure 5.2 and section 5.5 below). The employment of the same behaviours and morphologies on bifaces of different sizes by Gouzeaucourt knappers challenges McPherron's interpretation that what we see at GZnivG was solely the natural result of resharpening over time. This pattern indicates that the handaxes can be associated strongly in terms of a shared behavioural repertoire, something which is further strengthened by the same shape pattern occurring in several other levels at the Gouzeaucourt locality. As I have proposed in previous chapters, this consistent pattern of behaviour provides a strong foundation to investigate differences in skill-level. GZnivG hominins were doing the same things when making and curating handaxes, so one can compare bifaces within the assemblage to examine how skilfully one biface is made relative to another one.

Indeed, as part of my experimental investigation will explore, inexperience likely affected the fidelity with which different skill-levels could reproduce a particular form. Therefore, slight divergences from the mean morphology within an assemblage may be potentially correlated with other observable indicators of inexperience (e.g., increased failed removals culminating in step/hinge fractures, evidence of battering, and other irregularities and makers of poor performance). Certainly, consideration and examination of relative skill-levels of those responsible for creating bifaces at the site within a proposed shared social tradition of handaxe manufacture is something that needs to be explored further.

In the next section, I detail my approach to analysing morphological similarity at Gouzeaucourt, and the outcomes of that analysis. This was done to test previous cases of

morphological homogeneity (Field, 2002; McPherron, 1994; Soriano, 2000) and are entirely separate from results regarding my investigation of skill at GZnivG (and the experimental assemblage sample) which will be presented in Chapter 7. They are outlined purely to further demonstrate the suitability of the GZnivG handaxe assemblage for investigating skill-level differences in an assemblage of closely associated handaxes.

5.5 Morphological homogeneity at Gouzeaucourt

The sample used for this test of morphological uniformity at Gouzeaucourt Level G comprised the full assemblage of 161 bifaces from this level housed at the Maison de l'Archaeologie in Arras, France. Comparisons of handaxe shape regularity between Level G and other Gouzeaucourt levels were conducted, using samples from Levels H (100 bifaces) and R (52 bifaces). Levels H and R were assessed using photographs kindly shared with me by Dr. Lamotte, Université de Lille. Other comparative sites were Boxgrove, West Sussex (161 bifaces), due to the well attested high morphological homogeneity at the site, as well as the Broom Pits on the Dorset/Devon border (160 bifaces), known better for its morphological heterogeneity owing to it being a palimpsest assemblage recovered from gravel pits. Boxgrove and Broom Pits data was captured from 2D planform images of the handaxes stored on the Archaeology Data Service biface database (Marshall et al., 2002).

The comparison of GZnivG with other levels in Gouzeaucourt was to show how invariable the Gouzeaucourt morphology was over time. Comparison with Boxgrove and Broom Pits assemblages was to show how relatively homogenous the Gouzeaucourt level

G bifaces were, and whether they were more tightly similar like Boxgrove or more variable like Broom Pits examples. Unfortunately, due to the Covid-19 pandemic, returning to France to take photographs of CLG bifaces for inclusion here was not possible. This would have been ideal to show how heterogenous CLG is compared to the GZ levels given my arguments contra McPherron outlined above.

5.5.1. Rationale

McPherron's (2000, 1999, 1994) and Field's (2005, 2002) morphological analysis of homogeneity among Gouzeaucourt handaxes used dimensional proxies for shape, evaluating ratios of breadth, length and thickness first introduced by Roe (1968). However, their analyses did not account for the effect size differences had on the shape data produced among the assemblage. Indeed, allometry was something Crompton and Gowlett (1993) showed to be a major factor affecting Acheulean handaxe shape variation (Lycett et al., 2006). My approach will account for allometric differences in bifaces at Gouzeaucourt to test the constancy of morphology regardless of overall size.

5.5.2. Approach

In line with García-Medrano et al.'s (2020) WEAP method, I performed a Generalized Procrustes Analysis on the 2D planform morphologies of Gouzeaucourt handaxes from photographs (see Figure 5.4). To do this, I generated XY coordinate data at regular intervals (every 5% of the length) for points around the outline of the biface using 'AScheck', image-processing software developed by my colleague Stijn Debackere

(2020 https://github.com/StijnDebackere/AScheck) (TPSDig was not available to me). This data was scaled, rotated, and translated to bring all handaxe outlines to a standardised size, orientation, and position before subsequent analysis. This method worked to remove any allometric effects on shape analysis. It did this by size-adjusting the widths between the two points on the left and right of the biface at a given point along the length of the piece by using the geometric mean (García-Medrano et al., 2019; Lycett et al., 2006; Lycett and Von Cramon-Taubadel, 2008; McNabb, 2017). This allowed the assessment of differences in landmark shape data to be solely attributed to shape differences between separate objects and not their disparities in size. Using this data, the variance in biface shapes could be analysed at Gouzeaucourt.

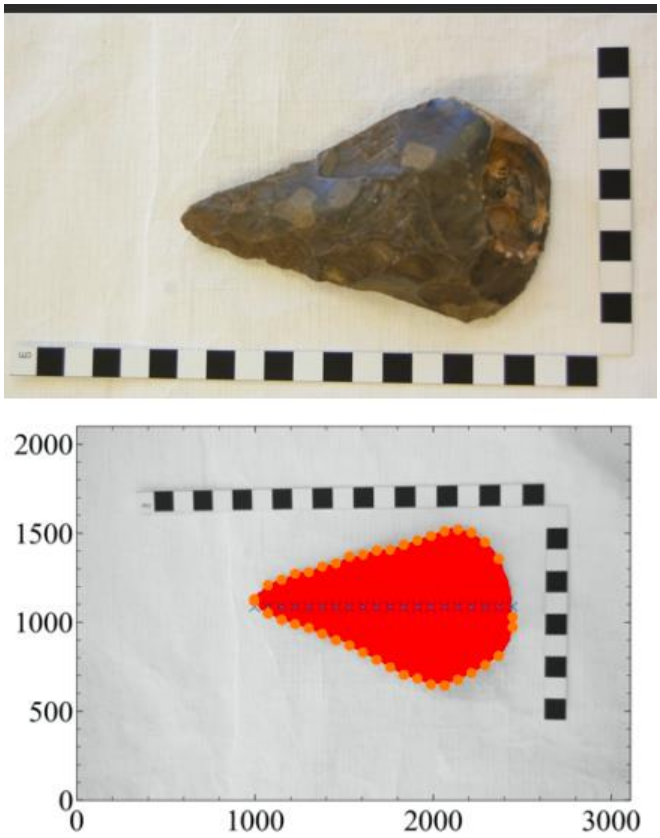


Figure 5. 4. Illustration of XY coordinate data captured through the AScheck method.

The size-adjusted, standardised shape data for the Gouzeaucourt Level G assemblage was then compared with similarly size-adjusted shape data from other assemblages and sites (Gouzeaucout Levels H and R, Boxgrove and Broom Pits). The shape data from these sites were subjected to a Principal Components analysis (PCA) in *PAST 4.0* software to assess the relative homogeneity of each assemblage in terms of planform morphology. Using the thin-plate splines functionality on *PAST 4.0*, shape data was analysed through visualising contour changes from the group mean of assemblages along a relative warp (García-Medrano et al., 2020). A Discriminant Analysis was also conducted (based on PCA results of 2D planform morphological variation) to test the power of shape variance to discriminate handaxe assemblages.

5.5.3. Evaluation

The first seven principal component scores captured 89.21% of shape variance among the sample analysed, with the first two principal component scores capturing 49.98% of variance among the sample from Gouzeaucourt Level G, Gouzeaucourt Level H, Gouzeaucourt Level R, Boxgrove and Broom Pits (PC1 = 29.42%, PC2 = 20.56%). This is a strong result suggesting there are individual influential factors affecting morphological variance that are captured in the dataset.

Figure 5.5 shows the scatter graph based on principal components 1 (PC1 - horizontal axis) and 2 (PC2 - vertical axis). Through the thin-plate splines functionality on *PAST 4.0*, PC1 displayed the range of more rounded morphologies below zero to more cordiform forms above zero with wider bases and more pronounced tips. Whereas PC2

showed the range from more exaggeratedly pointed forms signified by negative scores and more ovate shapes represented by positive scores.

In Figure 5.5, Gouzeaucourt Levels G (red) and R (grey) are visibly more rounded along the PC1 axis than Boxgrove and Level H and ranged evenly from ovate to cordiform along the PC2 axis. They are noticeably similar in terms of shape. Gouzeaucourt Level H (green) is more like the pattern of Boxgrove (black) along both axes and is focused more on the intersection of the two axes – being more positive in PC1 and PC2, suggesting a more cordiform shape than the other Gouzeaucourt levels but still being generally considered ovate.

Broom Pits (blue), on the other hand, is concentrated to the positive side of PC1 and is mixed regarding PC2, showing approximately half of the specimens to be like Boxgrove cordiform shapes and the other half being more distinctly pointed by the negative distribution of many specimens along PC2's axis. This highlights the palimpsest nature of the Broom Pits assemblage, like CLG outlined above.

The similarity of the Gouzeaucourt Levels G and R is striking. Some Gouzeaucourt bifaces, like the biface shown in Figure 5.1, have been subjected to post-depositional action and present as outliers in Figure 5.5. This affected the tightness of the cluster of GZnivG and GZnivR bifaces along both PC1 and PC2 compared to Boxgrove. However, the majority of GZnivG and GZnivR bifaces occupied a space between 0 and -2 along PC1 and 1 and -1 along PC2. Unsurprisingly, though, both levels G and R overlapped more with GZnivH than they did with Boxgrove or Broom Pits.

The greater PC1 ranges of GZ Level G and R bifaces than other samples was likely due to resharpening and later modification (see Figure 5.1), however, as mentioned above, the thin-plate spline functionality shows that the horizontal axis dealt more with variations on an ovate template, from rounded to cordate. Conversely, the PC2 axis more clearly covered the variation between ovates and pointed bifaces. On this axis (PC2), GZ bifaces exhibited a range closer to that of Boxgrove, and Broom Pits was by far the most variable. Almost half of Broom Pits' bifaces on the PC2 axis scored very negatively and did not correspond with scores of other sites closely, whereas the other Broom Pits bifaces intersected with Boxgrove and GZnivH. Arguably, this pattern reflects the reality that GZ and Boxgrove bifaces were made consistently ovate, whereas Broom Pits bifaces were comprised of a mixture of two templates of handaxe shape – ovate and pointed.

Therefore, the thin-plate splines functionality on *PAST 4.0* has revealed that PC1 and PC2 illustrate the broad boundaries of handaxe traditions. However, my analysis has also highlighted that intra-assemblage similarity can be extended between assemblages that are closely associated in time, space, and/or tradition (i.e., GZnivG and GZnivR). Nevertheless, individual assemblages can still clearly be differentiated from other assemblages belonging to the same tradition (i.e., GZnivG and GZnivR are very similar compared to Boxgrove). To further demonstrate this, I conducted a Discriminant analysis.

The Discriminant Analysis results presented in Table 5.1 helped to confirm the closeness of the planform morphologies (i.e., the shape of handaxes in frontal view – not from the side) of Gouzeaucourt in comparison to a well-known homogenous and closely related assemblage in Boxgrove and a well-known heterogenous and poorly related assemblage in Broom Pits.

	GZnivG	BXG	BROOM	GZnivH	GZnivR
GZ	58%	2%	1%	10%	29%
BXG	0%	52%	24%	25%	0%
BROOM	0%	33%	53%	14%	0%
GZnivH	4%	17%	14%	61%	4%
GZnivR	37%	4%	2%	12%	46%

Table 5. 1. Discriminant Analysis of biface shape variance at Gouzeaucourt (Levels G, H and R), Boxgrove, and Broom Pits, showing percentage of correct classification by assemblage.

The analysis showed a 54.57% accurate classification amongst the entire sample, with each assemblage being around the 50% mark for accurate classification. However, this obfuscated identifiable patterns of homogeneity at Gouzeaucourt. As seen in Figures 5.5, 5.6, and 5.7, Gouzeaucourt levels G and R were markedly similar, something that was borne out in the Discriminant Analysis results (see Table 5.1). Gouzeaucourt Level G was accurately classified to itself at 58% but, in addition, a further 39% was classified to other Gouzeaucourt levels – with only 3% being classified to non-Gouzeaucourt levels. In addition, Gouzeaucourt level R was accurately classified to itself at 46% but was additionally classified to other Gouzeaucourt levels at 49% - bringing the total classified to Gouzeaucourt at 94%. Although Gouzeaucourt level H was less markedly classified to Gouzeaucourt as a whole (total = 69%), it suggested there was distinct homogeneity across the site and within each level's assemblage.

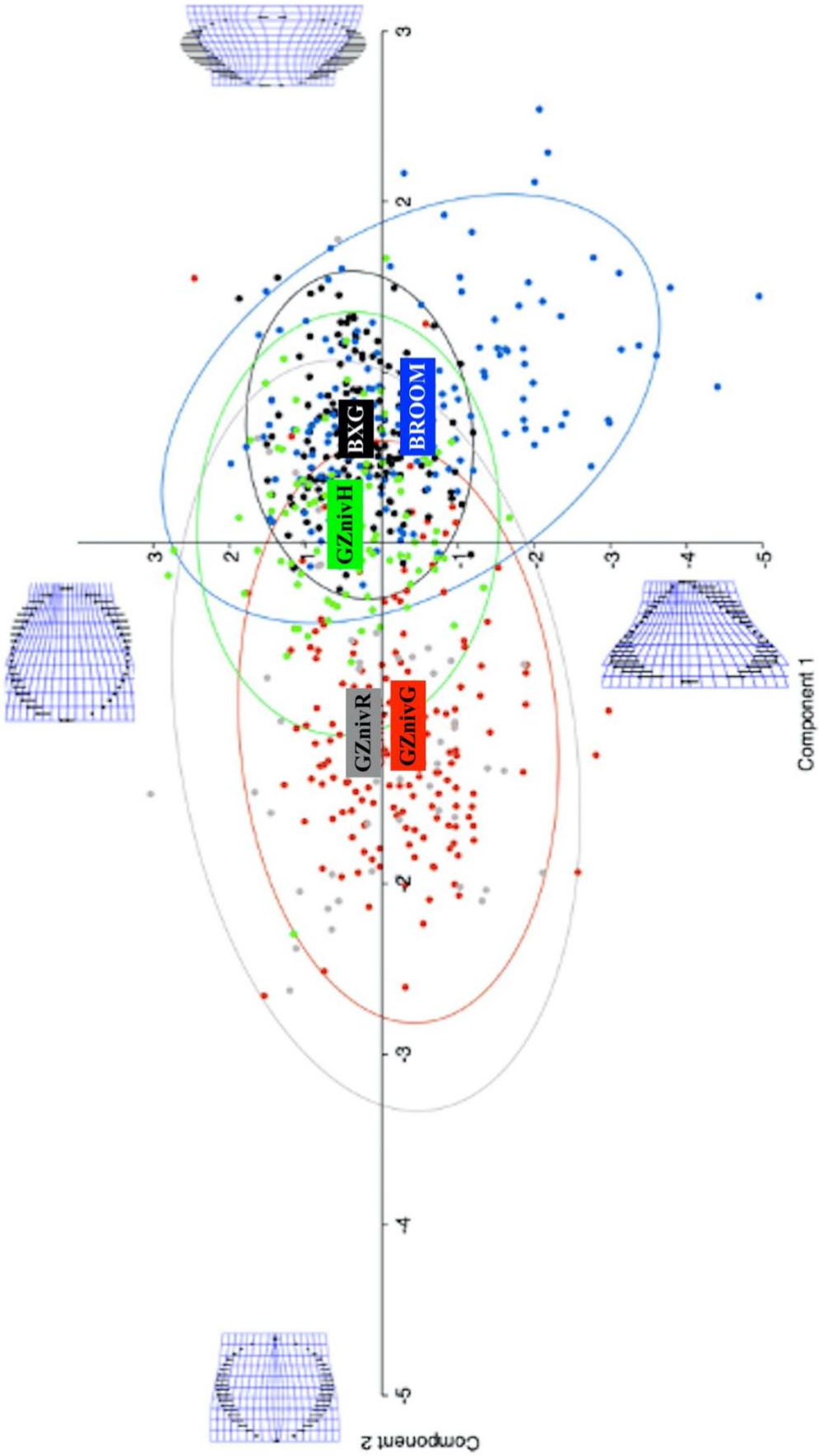


Figure 5. 5. Principal component scatter graph showing shape variance at Gouzeaucourt (G, H and R) Boxgrove and Broom Pits assemblages.

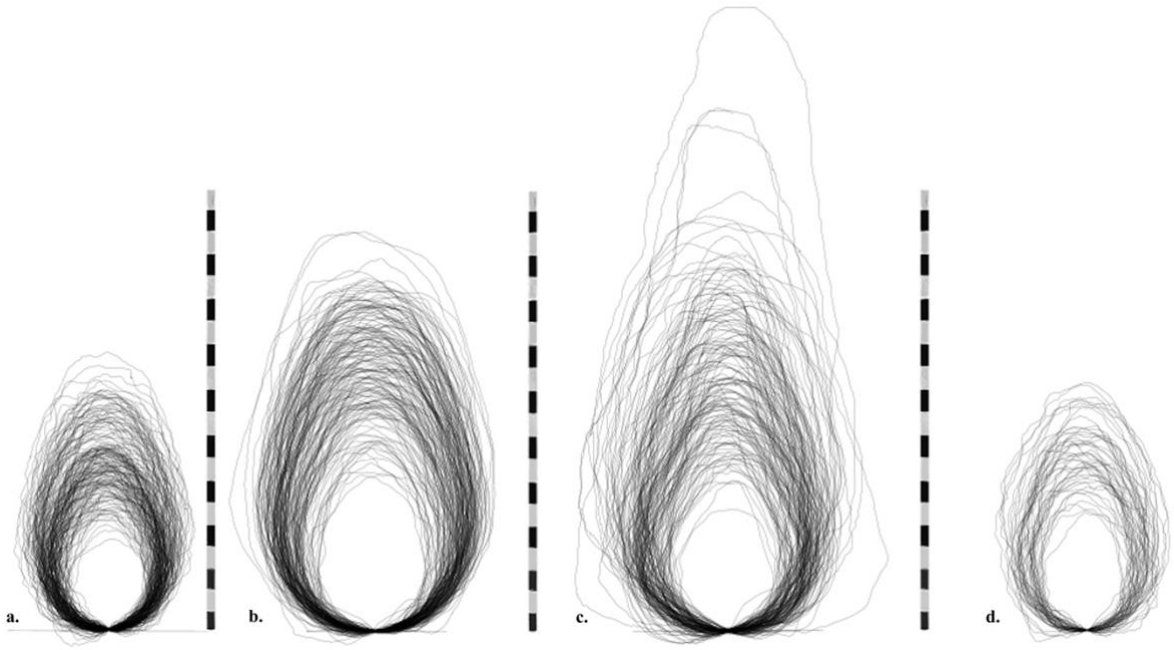


Figure 5. 6. Outlines of all bifaces from samples from GZnivG (a), Boxgrove (b), Broom Pits (C), and GZnivR (d). (Sources: Marshall et al., 2002 ADS biface database & Dr. A. Lamotte).

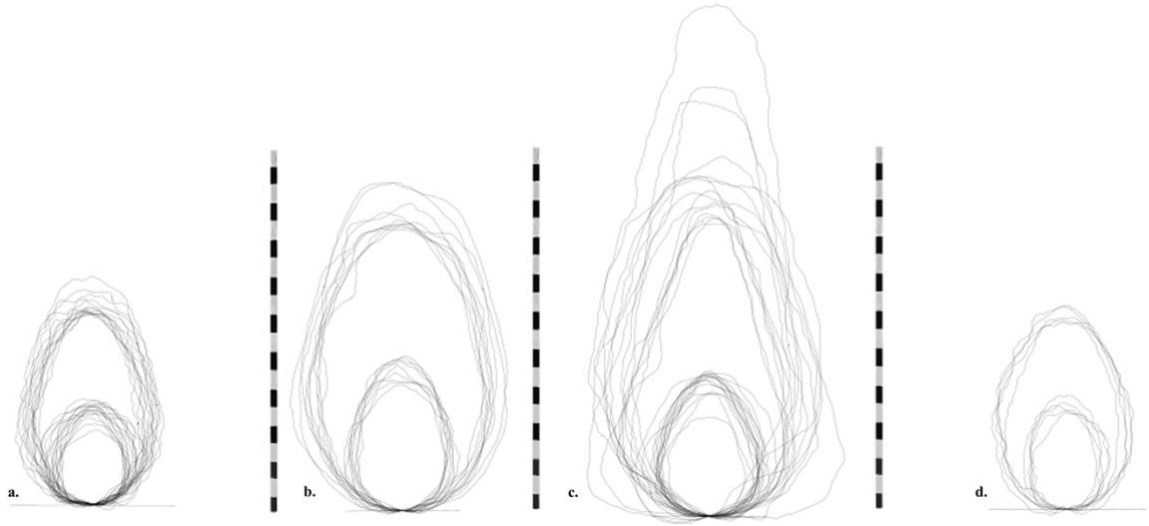


Figure 5. 7. Outlines of longest and shortest 10% of biface in samples from GZnivG (a), Boxgrove (b), Broom Pits (C), and GZnivR (d). (Sources: Marshall et al., 2002 ADS biface database & Dr. A. Lamotte).

I am confident that these results show a marked similarity within level G and across the other levels at Gouzeaucourt because I have used a large sample size that included a range of handaxe lengths. Thus, there being a general similarity in the shapes once their size

has been adjusted for, shows that handaxe size was not a major factor in changing the shape more generally.

Conversely, the pattern from Broom Pits was suggestive of a more palimpsest-like nature, as mentioned above, with 47% classified to Boxgrove and Gouzeaucourt Level H (more ovate and cordiform morphologies) and 53% classified to itself which more strikingly represent pointed morphologies (see Figures 5.6 and 5.7). As Figures 5.6 and 5.7 show, pointed forms at Broom Pits are of various sizes, thus providing more evidence that planform shapes were intentional and not solely the result of resharpening from large points to small ovates. In addition, the fact that neither Broom Pits nor Boxgrove had any specimens that were classified for both Gouzeaucourt Level G and Gouzeaucourt Level R is significant, suggesting (in addition to the homogeneity between Levels G & R) that Levels G and R represent a spatio-temporally removed, culturally distinct tradition of biface manufacture, which included resharpening.

5.6 Summary

The rationale presented above provides a justification for how I will, in this thesis, investigate skill-level differences between Acheulean handaxe makers. Key to my reasoning was explaining the importance of experimental investigation of modern day knappers in providing the necessary data for analysing skill-level differences in the archaeological record. Equally important to my rationale was outlining the contextual parameters for investigating skill-level differences in the Acheulean. I have argued that investigating these differences is only relevant between handaxes in an assemblage that

shared a coherent goal of manufacture. I, therefore, provided an overview of prior scholarly debates, as well as my own analysis of biface shape, to justify utilising GZnivG to investigate differences in skill-level in the Acheulean. I must reiterate that I will not be investigating the differences in skill between individuals but categories of individuals – classified by their skill-level.

In the next chapter (6), I outline the quantitative and qualitative methods and analyses performed on both experimental and archaeological assemblages for the purpose of investigating skill at Gouzeaucourt Level G.

6. Chapter 6 - Methodology

6.1 Introduction

In the previous chapter I demonstrated the suitability of using GZnivG to investigate differences in skill-level in the Acheulean record by showing that knappers there behaved very similarly during biface manufacture. Not only did Gouzeaucourt knappers produce extremely consistent planform shapes but they employed the same knapping techniques as well (Field, 2002; McPherron, 1994; Soriano, 2000; Tuffreau and Bouchet, 1985). This was an important baseline to establish and strongly suggests that the knappers were linked together in a shared pursuit, permitting comparison of bifaces based on how well they adhered to it. However, because the skill-levels of the GZnivG knappers were undetermined, discriminating between them properly required using insights from an assemblage where the skill-levels of the biface-makers was indisputable (i.e., among modern knappers). Thus, observations and data gathered from my experimental investigation will be used directly to generate a framework for differentiating between bifaces by skill-level category.

Nevertheless, the boundaries between skill-levels in contemporary contexts can sometimes be ill-defined. Despite acknowledgement by knappers and researchers alike that knapping skill increases along a spectrum from poorly skilled to highly skilled, a methodological lacuna exists in breaking the binary distinction of Expert and Novice and identifying those who are skilled but are not masters (i.e., Intermediates). Intermediates, or moderately skilled individuals, are categorised by holding a few years of experience

but not the prolonged practice of Expert knappers. Indeed, they are competent enough to make functional Acheulean bifaces regularly and without the frequent trouble that beginners encounter. Notwithstanding this, they are unable to attain the performance levels of master knappers, and do not possess Expert-level flexibility to overcome more difficult problems or mistakes. Therefore, as will be outlined in this chapter, I included Novice, Intermediate, and Expert knappers in my experimental investigation to test whether an additional degree of skill discrimination can be achieved and applied to the analysis of archaeological assemblages.

6.2 Methodological approach

To differentiate between bifaces, my methodological approach needed to refrain from being too narrow, focused, and vulnerable to anomalies. Consequently, I adopted a mixed methods approach, utilising both quantitative and qualitative methods of analysis. This allowed my results to be compared, integrated, and interpreted more broadly. It also afforded scrutinisation of both methods by simultaneously ascertaining their limitations (Bouchrika, 2021; Brannen and Moss, 2012; Haq, 2014).

Identical criteria (quantitative and qualitative) were used to analyse both experimental and archaeological biface assemblages. Realisations gained from my experimental study helped inform my interpretation of handaxes from GZnivG. As I explain in more detail below, by finding the quantitative thresholds for known experimental skill categories and transposing them onto the archaeological data, I was able to generate skill-level classifications for each GZnivG biface. However, among both experimental and

archaeological assemblages, misclassifications were possible due to occasional mistakes, differences in raw material size and/or quality, disparities in the morphologies attempted, as well as instances of lesser-skilled participants making bifaces under expert instruction.

To evaluate the reliability of my quantitative classifications of both experimental and GZnivG bifaces, they were compared against the results of previous qualitative assessments I had made during initial data collection. This provided context for elucidating why a misclassification of proficiency may have occurred. In addition, to further evaluate the efficacy of my quantitative criteria, I applied the experimentally derived skill-category thresholds to a Blind Test of one Expert and one Intermediate. Moreover, a questionnaire, granted Ethics approval by the University of Oxford, combined a mixture of quantifiable questions and broader discussion-based questions to shed further light on the results and aid discussion of skill-level more generally.

6.3 Materials

Before detailing the quantitative and qualitative data collected and the analyses performed, an overview of the materials used in both experimental and archaeological investigations is provided.

6.3.1 Experimental assemblage

The experimental investigation involved the participation of 7 variably skilled contemporary knappers. The participants were sorted into three skill-level categories: Novice, Intermediate, and Expert. Skill-level classification was based on the number of years of practice they held. Novices were classified as those with less than one years' experience, Intermediates were categorised as possessing between one- and fifteen-years' training, and Experts retained over 15 years of engagement in knapping activities.

The rationale behind the classification boundary between Expert and Intermediate was broadly based on the 10,000-hour rule (Gladwell, 2008). 10,000 hours divided by the number of days in fifteen years, equates to 1.8 hours per day. This is not proposed as an actual daily schedule for skill development, but it is not an overly ambitious estimate. All Experts that contributed to this research had considerably more than fifteen years' experience, and both Intermediates had under 5 years' practice. Therefore, I consider the boundaries used to be conservative and reasonable. The reasoning behind the classification boundary between Intermediate and Novice was based on Pargeter et al.'s (2019) finding that the manual dexterity needed to reliably detach large thinning flakes indicative of more skilled performance required circa 200 hours of practice. Over a year, 200 hours of practice corresponds to around 33 minutes per day on average. I therefore regarded 1 year to be a realistic timeframe for 200 hours to be accumulated by a learner and is accordingly an appropriate threshold to separate knappers in my experimental investigation.

My inquiry was designed to gather as close to equal samples as possible from each of the three skill-levels. However, the Covid-19 pandemic interrupted in-person data collection. Fortunately, I was able to increase my sample size of handaxes for all skill-level

categories through gathering data from the University of Liverpool's experimental collections. Thus, not all experimental bifaces were manufactured under my direct observation. Nevertheless, I managed to witness multiple knapping sessions with all skill categories. For both Novices and Intermediates, I observed them knapping alone and under instruction from an Expert. Similarly, I observed Experts knapping in isolation as well as knapping while giving instruction to Intermediates and Novices. My observations were used in conjunction with results of a questionnaire (outlined in more detail in section 6.3.3) given to other knappers, not all of whom could take part in the experimental investigation. Ultimately, this supplied additional anecdotal evidence to the more generalised theoretical approaches to skill presented in Chapter 4. Used together, they permitted a more comprehensive discussion of competence to take place in relation to biface manufacture.

The experimental assemblage was comprised of 130 handaxes made by 7 individuals (participant identification numbers and their handaxes are presented in Catalogue 4):

- Novice – 47 (3 individuals: Experimental knappers 4, 6, and 7)
- Intermediate – 20 (2 individuals: Experimental knappers 4 and 5)
- Intermediate Draft – 25 (1 Individual: Experimental knapper 4)
    - Deemed Draft as the handaxes were discarded as unfinished or unusable (due to many mistakes).
- Expert – 38 (3 individuals: Experimental knappers 1, 2, and 3)
    - One Expert only contributed two handaxes. The remaining 36 were made by the other two Experts.

It should be noted that I, myself (Experimental knapper 7), contributed to the experimental data collection. Excluding the Novice bifaces that were sourced from the University of Liverpool collections, my contribution was made over a period of 6 months, whereas the other Novice (Experimental knapper 6) made a one-time contribution. As a beginner, I observed how other skill-levels dealt with problems and linked that back to the challenges I faced in becoming more competent. This served a dual purpose. Firstly, my involvement generated quantitative data related to how practice affected skill development. I could also compare this to the degree of influence (quantitative) that Expert instruction had on the handaxes I produced. Secondly, it furnished me with greater qualitative understanding regarding how one's proficiency progressed in relation to discernible aspects of the handaxes created. Relating my insights to the theoretical approaches to skill outlined in Chapter 4, aided my discussion of what skill is and how it improves.

I should also draw attention to the fact that one Intermediate (Experimental knapper 4) regularly made handaxes for their own research and possessed an extensive collection covering the years of practice in which they had made each handaxe. Hence, this participant also contributed 10 bifaces, made in their first year of practice, to the Novice group. This participant was also responsible for the Intermediate Draft group. Data from 'Draft' bifaces was only collected from Experimental knapper 4 because no other knapper had kept a catalogue of bifaces they deemed unfinished or unusable. Therefore, although a 'Draft' subcategory would have been an interesting dataset to obtain from all skill-levels, I was limited by the material available to me. Nevertheless, this 'Draft' dataset was included in my overall analysis because, as questionnaire participant 7 noted, some bifaces are rejected as incomplete or unusable by the knapper in the Acheulean. Thus, I

reasoned that ‘Draft’ bifaces would provide useful supplementary data to the majority ‘completed’ bifaces from each of the skill-levels.

No pre-set constraints were imposed on any of the knappers. The bifaces that were made under my observation were simply the result of knappers being instructed to make an Acheulean-style handaxe. Likewise, the bifaces that formed part of experimental collections at the University of Liverpool were also made to no specific morphological parameters. Reducing morphological bias and constraints on making bifaces in the experimental sample was important for determining how shape and method of manufacture (i.e., reduction of a nodule or production on a flake) affected skill-level classifications, as will be elaborated on below.

Lastly, a Blind Test was devised to assess the efficacy of my methodology to accurately classify bifaces based on their quantitative criteria (see Catalogue 5). The Blind Test comprised 6 bifaces, 3 made by an Intermediate (Experimental knapper 4), the other 3 made by an Expert (Experimental knapper 1). Since both participants had contributed handaxes to the main experimental assemblage, it was necessary to impose constraints on manufacture. The aim of this was to reduce the effect of any morphological preference bias that each knapper may have held and I may have subconsciously, or consciously, attributed to an individual. Both contributors attempted to replicate an experimentally made biface specimen retained by the University of Liverpool. The biface, shown in Figure 6.1 below, is highly refined and made by a master knapper with multiple decades of experience (John Lord).

The Blind Test sample size was not as large as I would have liked, but due to the complications regarding Covid-19, it was all I was able to organise at the time.



Figure 6. 1. Blind Test example biface made by John Lord.

6.3.2 Archaeological assemblage

As discussed previously, the morphological and technological consistency exhibited by the Gouzeaucourt knappers can be used as a control variable for investigating skill-level. The GZnivG assemblage initially consisted of 161 specimens to be included in my analyses. However, several factors conspired to reduce this to 129 bifaces, a similarly sized sample to the experimental assemblage. As briefly explained in Chapter 5, the sample was reduced due to some bifaces exhibiting extensive secondary working (see example in Figure 5.1). This affected the morphology in some cases and, as such, negatively impacted the efficacy of the quantitative criteria used to discriminate between skill-level categories. Secondly, two key skill-related quantitative criteria (Edge straightness and Profile Asymmetry) were not collected for some bifaces due to the cancellation of a data collection trip to France because of the pandemic.

Fortunately, full data collection was completed for a total of 129 bifaces. Additionally, all initial qualitative observations regarding skill-level were made over the two data collection trips I had previously made. However, diacritical analysis, which I had hoped would be more of a central feature of my research, had to be amended to an alternative method of performing it via photographs. This was a sub-optimal solution, as diacritical analysis typically benefits more from the analyser physically observing the biface in three dimensions.

6.3.3 Questionnaire

Following discussion with my Transfer of Status assessors, I devised a questionnaire aimed at obtaining testimony from contemporary flintknappers regarding their own skill

development, what aspects of biface manufacture they regarded as the most challenging, as well as their more general thoughts on accomplished performance.

Primarily, the questionnaire was designed to hold a mirror to the quantitative and qualitative data collection that formed my mixed methodological approach. It presented an ethnographic lens through which to contextualise my theoretical and methodological framework and discuss the suitability of my criteria for investigating skill-level differences of handaxe-makers in the Acheulean. Further insights generated by questionnaire contributors were applied to the construal of, particularly experimental, results, in addition to the viewpoints given by experimental participants. Eight individuals from the UK, France, and the US partook, all of which are interested in the topic of skill in the Acheulean context and most of whom have knapped or consider themselves knappers.

6.4 Data collection

Data collection was organised with the analyses I wanted to perform in mind, for both experimental and archaeological assemblages. These are set out in section 6.5.

Identical instruments were used for quantitative data collection. Callipers and scales were used to gather dimensional data which were used to calculate skill-related criteria indices such as Ratio of refinement. Photographs also permitted further dimensional data to be captured. When inputted into the AS-check program, introduced in Chapter 5, biface

asymmetries could be calculated. Pixel data was also taken from photographs to compute Edge straightness.

Qualitative data was likewise collected for both experimental and archaeological assemblages in mostly the same way. Instances of skilful performance, such as tranchet sharpening and the removal of extensive thinning flakes, as well as irregularities and mistakes not measured in my quantitative analysis, were recorded. In addition, diacritical analysis as per Lamotte (2012, 2011) provided supplementary data to aid qualitative descriptions of each biface from GZnivG (see Catalogue 1). Another source of information regarding skill in relation to handaxe manufacture was generated via my questionnaire of current knappers.

6.4.1 Quantitative data

The quantitative data collected is itemised below, alongside rationale for their collection.

Dimensional data (*Length (mm), Breadth (mm), Thickness (mm), and Weight (g)*)

- Collected using digital callipers and scales and combined to detail the range of sizes in both biface assemblages (experimental and archaeological). Although not relevant for determining skill, they also provided information on the variation of metrical data for known skill-levels within the experimental dataset.

Breadth/Length ratio

- Calculated by dividing biface breadth by its length. It has previously been used to comment on the aesthetic preferences of hominins (Pope et al., 2006).

Alternatively, Gowlett (2011) demonstrated that it more directly related to the length of the biface. He found that as biface length decreased, Breadth/Length ratio increased. Hutchence and Debackere (2018) have corroborated Gowlett's (2011) findings. However, I wanted to test whether this could be replicated in an experimental assemblage, and if skill-level affected the pattern observed.

Roe's (1968) L1/L determination of biface shape

- Calculated by finding the distance of the point of maximum breadth to the base of the biface and then dividing it by the length of the biface. This was Roe's (1968) quantitative method for determining pointed, ovate, or cleaver shapes. The thresholds for metrical points were those bifaces with indices below 0.35, ovate bifaces exhibited a range from 0.35 to 0.549, and those with scores of 0.55 or more were nominated as metrical cleavers, having their maximum breadth nearer to the tip than the butt of the biface.
- L1/L was used to assess whether the skill-criteria used to discriminate between skill-levels in my experimental analysis were affected by biface shape. No morphological constraints were imposed on experimental bifaces and, as such, the assemblage consisted of both ovates and points. However, the GZnivG bifaces are overwhelmingly ovate. Therefore, it was necessary to assess whether my criteria for discriminating between skill-levels among modern knappers was more accurate for ovates than points, or similarly accurate for both. The outcome of this informed which results were used to interrogate the archaeological results from GZnivG.

Ratio of Refinement

- Calculated by dividing the thickness of the biface by its breadth, it generates an index from 0 to 1. Scores closer to zero indicate greater refinement. Some authors (e.g., Stout et al., 2014) have reversed the ratio and divided breadth and/or length by thickness. However, this creates indices between 1 and 100, which I found unhelpful when comparing to other quantitative measurements with indices between 0 and 1.
- 'Refinement' refers to the ability of the knapper to thin the biface effectively without reducing the breadth of the piece. Usually, breadth reduction is caused by inaccurate edge maintenance, platform preparation, striking platform selection, and flake striking accuracy. Therefore, it is a method for quantitatively assessing the thinning abilities of the knapper.
- However, as mentioned in Chapter 5, McPherron (1999) identified that beyond a certain threshold of reduction/resharpening, thickness increases relative to breadth and length at Gouzeacourt as bifaces become smaller. Therefore, significantly re-sharpened and very small bifaces from GZnivG could potentially be negatively affected by using refinement ratios in their quantitative skill-level classification.

Thickness/Length ratio

- Calculated by dividing the thickness of the biface by its length and, like Ratio of Refinement, it measures the relative ability of the knapper to keep the dimensions of the biface elongated while thinning the biface effectively for functional cutting purposes.
- The same threshold (McPherron, 1999) outlined above for Ratio of Refinement also applies to Th/L ratio.

Planform Asymmetry

- Calculated in AS-check - a Python script developed by my colleague Stijn Debackere (https://github.com/StijnDebackere/AScheck; Hutchence and Debackere, 2018). AS-check, as outlined in Chapter 5, can also be used for calculating shape data using XY coordinates. However, its initial purpose was to compute asymmetry indices for objects in 2D photographs. AS-check does this by quantifying the number of asymmetrical pixels ('*Ap*') relative to total number of pixels ('*Tp'*) constituting the object, generating an index of asymmetry.
- Equation: ***Ap/Tp = asymmetry index (%)***
- Results above '1' are incorrect, indicating the failure of the script to discern the object from the background. Although similar to Hardaker and Dunn's (2005) 'Flip Test', the AS-check script crucially differs by permitting the asymmetry indices of multiple images to be calculated simultaneously.
- Other than symmetry being considered as intentionally imposed by knappers (McNabb and Cole, 2015; White and Foulds, 2018), there is no clear consensus regarding whether it was more functionally or socially significant. Machin et al. (2007) found a weak correlation between increasing planform symmetry and increasing butchery effectiveness. Yet, other, less functional explanations, such as Kohn and Mithen's (1999) sexual selection hypothesis, Machin (2009) linking it to 'status', or Pope et al.'s (2006) 'Golden ratio' argument, imbued biface symmetry with cognitive and social evolutionary significance.
- Broadly speaking, symmetry has often been linked with greater refinement and more skilful behaviour. Yet, Winton (2005) showed that inexperienced knappers can reach similar levels of planform symmetry to Experts. Despite this, Planform asymmetry is used here to test its discriminatory power to differentiate, when

combined with other criteria in a multivariate analysis, between bifaces of known-skill-level in the experimental investigation.

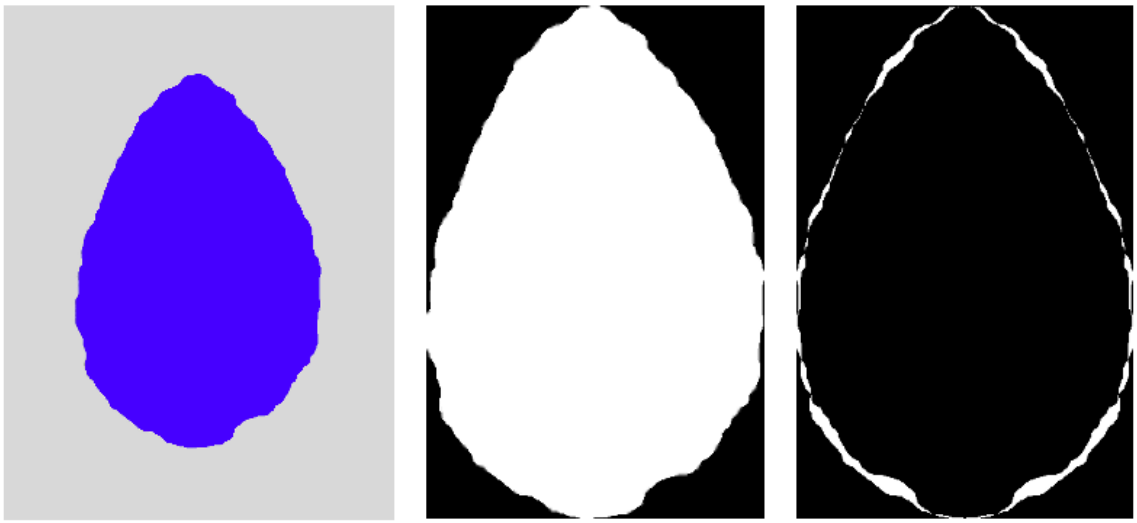


Figure 6. 2. AScheck process of calculating symmetry of GZnivG biface 2812(1). The original image is converted into grayscale and the biface is separated from its background. The black and white image is then used to find the planform contour. This contour is then flipped around the longest axis of the biface (assumed to be the axis of symmetry). The difference between minimum and maximum symmetric outlines results in the number of asymmetric pixels which can be used in the equation above.

Profile Asymmetry

- Calculated as per Planform Asymmetry.
- Machin et al. (2007) found no relation between this and butchery effectiveness. Though, creating symmetrical bifaces in profile is a more difficult skill than creating symmetrical bifaces in planform.
- Although hominins possessed an innate ability, like living humans, to recognise symmetry (Hodgson, 2015; Wynn and Gowlett, 2018), its manifestation in profile requires a greater appreciation of the biface manufacture process. To create a

symmetrical biface in profile, a knapper needs to control the plane of intersection where the two faces of the biface meet at the cutting edge. Indeed, Shipton (2019) noted that profile symmetry is inextricably linked to edge maintenance and thinning, where flaking of high topographical areas on either face needs to be managed so that the cutting edge does not become sinuous as a result. For this reason, it is a skill-related criterion in my analysis.

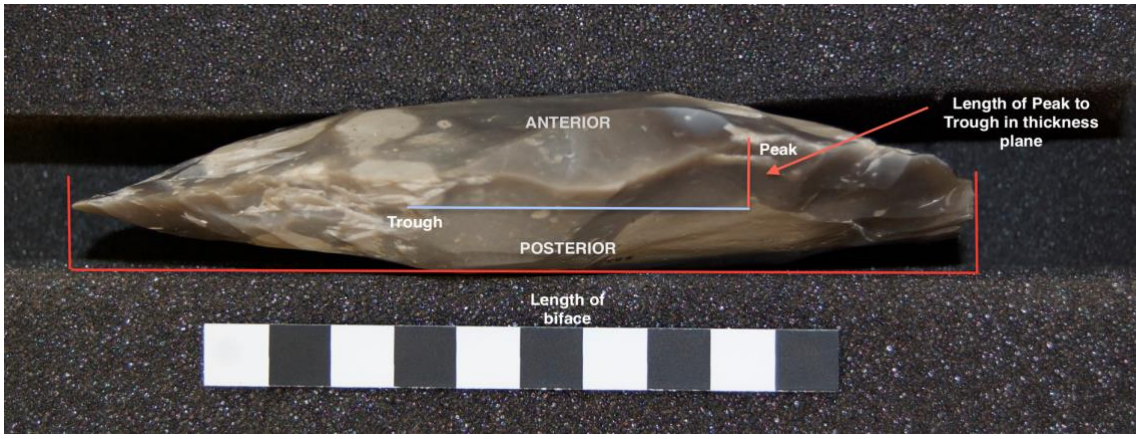


Figure 6. 3. Illustration of the measurements needed to calculate Edge straightness.

Edge Straightness

- Measured using photographs of bifaces in profile and borrowed from Shipton's (2018) method, which calculated the width of the peak and trough of the handaxe edge in the thickness plane (i.e., profile view) relative to its overall length. The peak and trough are calculated by viewing the biface in profile and deciding upon points along the edge that are furthest towards the anterior and posterior faces (see Figure 6.3). The length between these points (straightness) is measured in the thickness plane (Figure 6.3) by number of pixels to negate any inherent bias caused by differences in the levels of perspective in photographs.

- To generate an Edge Straightness index, the pixel length of the straightness measurement in the thickness plane is divided by the pixel length of the biface. This was done for both edges and the mean of both edges was used for the analysis I will present here.

Flake Scar number

- Inspired by data originally collected by Shipton and Clarkson (2015), it is determined by counting the number of flake scars larger than 10mm on either face. They found that the relationship between the volume of a biface and the number of flake scars above 10mm it exhibits is a reliable indicator of reduction intensity. Ultimately, they used this data to test whether reduction intensity affected biface shape, as per McPherron's re-sharpening hypothesis, and found that it did not. Instead, biface shape was conserved independently of reduction intensity.
- I was unable to measure the volume of the bifaces at GZnivG because I did not have access to the same scanning equipment used by Shipton and Clarkson (2015). However, I could measure Flake Scar numbers (>10mm).
- My observations of knappers of different abilities and experience levels suggested that this measurement could be used for another purpose (i.e., a greater number of flake scars may also indicate one's relative skill-level). Experts regularise, or retouch, their biface edges more, and through a series of processes of thinning and shaping, create multiple layers of flake scars on biface faces. Whereas less skilled knappers do not have such a controlled and methodical approach.
- Among later Acheulean bifaces, as Stout et al. (2014) noted, platform preparation became more common. In addition, Soriano (2000) noted that at Gouzeaucourt, edge regularisation was regularly performed. Thus, the number of flake removals

is a potentially relevant data source to examine. I sought to see if there was any discriminative power in this variable amongst my experimental assemblage sample.

6.4.2 Qualitative data

The qualitative data collected is detailed below, alongside the rationale for their collection.

Evidence of excessive force

- Incipient cones (presence/absence), crushing of edges, and endshock.
    - Indicate the use of excessive force by the knapper.
    - As Bril et al. (2010), Nonaka et al. (2010), and Williams-Hatala et al. (2020) have shown experimentally, Novices were less able to control the kinetic energy used in detaching flakes of varying sizes whilst utilising different weights of hammerstones.

Mistakes

- Failed flake terminations (step/hinge/snapped)
    - Linked to the use of too much force, incorrect striking angle, and poor platform preparation.
    - Multiple instances of failed removals indicate the knapper's inability to engage with the material successfully. Alternatively, they could signal raw material irregularities, such as inclusions.

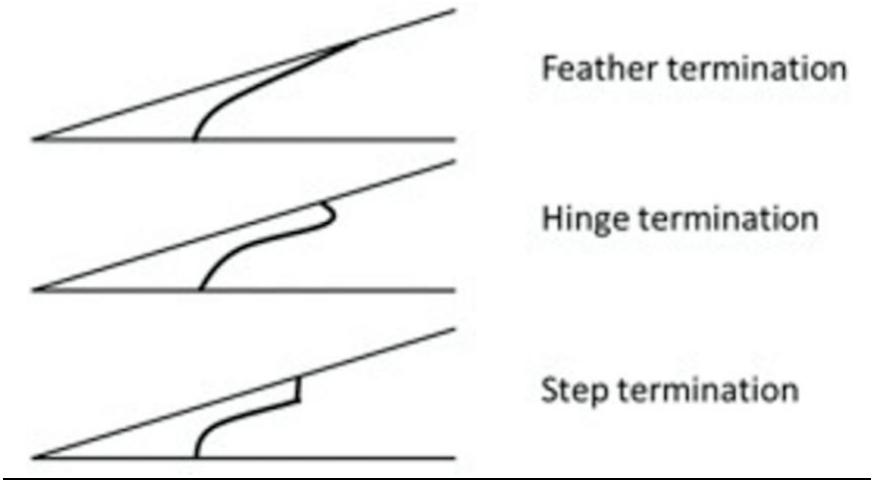


Figure 6. 4. Flake terminations. (Source: Coppe and Rots, 2017)

Raw material irregularities

- Inclusions
    - Can affect successful flaking, and sometimes lead to failed flake terminations. Qualitative assessment will assess how well the knapper worked with this hindrance to knapping. Moreover, failed flake terminations resulting from poor raw material or inclusions can affect quantitative scores. Hence, evaluation of the state of the material is key to judging whether quantitative skill-level classifications are fair.
- Initial raw material shape
    - Can limit the quantitative scores that can be achieved (e.g., if the initial nodule is small, round, and thick). If assessment of this is possible, it will also be able to qualify whether quantitative skill-level classification is fair or not.

Behaviours not captured by quantitative analysis

- Presence/absence of tranchet flake removal (at GZnivG)
    - It is a difficult skill and assessment of how successful it has been implemented may be able to clarify quantitative skill-level assessments which are borderline.
- Amount of cortex removed
    - This may affect quantitative classification, and assessment of the biface may give answers to why not all cortex has been removed and whether the quantitative assessment is fair.

6.4.3 Diacritical data

The diacritical data was only collected for the archaeological assemblage and followed Lamotte's (2012, 2011) method (see Figure 6.5) for determining the pattern of flake removals employed by the knapper. As mentioned above, I was unable to physically perform this in-person with the bifaces due to the COVID-19 pandemic. Lamotte's (2012, 2011) alternative approach allowed lower-resolution diacritical data to be captured from previously taken photographs. This information provided another quantitative way of analysing differences between the bifaces.

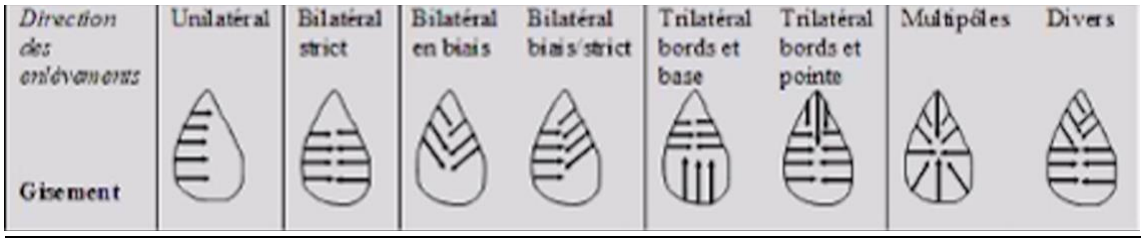


Figure 6. 5. Diacritical analysis as per Lamotte's (2012, 2011) method.

6.4.4 Questionnaire data

The responses to questions generated a mixture of quantitative and qualitative data. The quantitative data tested how differently modern knappers agreed or disagreed with certain statements regarding skill and biface manufacture. For these questions, participants were required to provide an answer on a scale of 1 to 5 regarding how strongly they agreed with the statement (1= Not very, 5= Very).

The qualitative data generated was in the form of unranked responses that required those who partook to elaborate on a series of questions related to knapping and their thoughts on skilled performance. The aim of these broader questions was to yield insights regarding the suitability of my analytical criteria, as well as what the contributors consider as skillful. Both sets of data were used to inform my discussion of skill more generally and my interpretation of the results more specifically.

6.5 Analyses

The analyses performed on the experimental and archaeological biface assemblages were similar but differed in a few ways. The reason for this mainly lay in the desired outcome of the analysis and the material being examined.

For example, both experimental and archaeological biface assemblages were subjected to univariate and multivariate statistical analysis, though, only the experimental assemblage

was exposed to bivariate correlations and scatter plot analysis by skill-level. The reason for this was because it was necessary to assess whether key skill-related criteria correlated in a statistically significant manner among bifaces with known skill-levels before using such criteria in the multivariate analysis of the archaeological assemblage. Moreover, the fact that the skill-level category associated with experimental bifaces was known, allowed me to perform analyses investigating how shape, as well as practice and instruction, impacted the ability of quantitative criteria to discriminate between bifaces made by knappers of varying experience. These analyses were not replicated with the archaeological assemblage because the skill-levels of the bifaces were undetermined, however, they were utilised in the shaping of the analysis of GZnivG. This will become clear in the Results chapter (Chapter 7).

6.5.1 Experimental assemblage

6.5.1.1 Quantitative analyses

The first quantitative analysis presented in my results were the univariate range, interquartile range, mean, and standard deviation statistics of each criterion outlined above by each skill-level. Presentation of univariate statistics provided a broad introductory overview of the quantitative differences between the skill-levels.

Via Spearman's rank *rs* bivariate correlations, I then tested whether skill-related criteria correlated with each other in a statistically significant manner. Correlations were undertaken in IBM SPSS Statistics (version 26.0). The objective was to test the results of Hutchence and Debackere's (2018) study which showed that skill-related criteria did

correlate statistically significantly. In addition, I wanted to assess whether indices indicative of greater skill were more likely to be correlated than not.

I performed Spearman's Rank, and not Pearson's, because, *a priori*, I did not expect linear correlations, but more generalised correlations. Spearman's Rank is a more general (and non-parametric) test for correlations, whereas Pearson's is used to test the correlation of parametric data. It is important to note here that bivariate correlations only look for common behaviour, and do not demonstrate a link between the two variables as a linear regression would do. It is possible that an external factor is controlling the correlation pattern of two variables.

Statistical significance was clearly highlighted and designated at both a 95% and 99% confidence level, in line with generally accepted significance thresholds. Strong correlations were indicated where $r_s > 0.5$. This is because, given the variance in correlations, scores lower than 0.5 do not necessarily identify strong correlations, even if the *p*-value is significant. To further examine whether the indices I proposed to be indicative of skill correlated and were reached by higher skill-levels in my experimental assemblage, I also produced a bivariate group scatter graph showing each biface classified by skill-level using a colour-coded system.

Next, I performed multiple multivariate discriminant analyses aimed at assessing the efficacy of the quantitative criteria for discriminating between skill-levels of experimental participants. Specifically, I used Linear Discriminant Analysis (LDA), a form of Discriminant Function Analysis (DFA). LDA distinguishes between known skill-levels by calculating the probability of group membership based on predictor variables

that have been calibrated on a set of measured variables (i.e., skill-related criteria) with recognised group labels (i.e., known skill-levels). This analysis looked at the relationship between the skill-related criteria by documented skill-level group and predicted whether the scores a biface achieved were 'truthful' to the group it had been assigned.

LDA assumes that different groups (i.e., known skill-levels) will possess different means for each measured variable, or skill-related criteria (e.g., Ratio of refinement). However, it supposes that for each criterion the independent groups will be identically distributed around their own mean. A critical part of the LDA assumption is that each measured variable has identical covariance within separate groups. Yet, if the different groups' distribution in each criterion/variable does not have the same covariance, a separation can still be performed if the measured values for the individual groups differ enough (i.e., if the means of the separate groups differ by more than both groups' standard deviation/variance).

To guard against overestimating the predictive power of the LDA, jackknifing was performed. This works by excluding one individual biface in the sample and generating an axis using the remaining data. This axis is then used to categorise the biface removed from the analysis, meaning a self-fulfilling classification, that uses the biface's own data to classify itself, is avoided. The accuracy of the resulting classifications was assessed relative to the known skill-level of the knapper who made them. This provided a percentage accuracy score which related to the effectiveness of the quantitative criteria used. This method was used to conduct multiple LDAs consisting of different combinations of criteria to find the most accurate combination for correctly grouping

handaxes by their known skill-level. Additional detail on LDAs and how they are reported are included in Appendix 1 and in sub-section 6.5.2.1 below.

As outlined briefly in Chapter 4, I will use these methods as a predictor of the following possible outcomes:

Hypothesis 1 (H^1):

- Each skill level (Expert, Intermediate, and Novice) will be quantitatively distinguishable from each other.

Hypothesis 2 (H^2):

- Quantitative misclassification will be most prevalent among Intermediate- and Novice-made handaxes, and Experts will rarely be quantitatively misclassified due to their multiple years more experience.

Hypothesis 3 (H^3):

- Expert knappers will generate a tighter range of quantitative scores than Intermediates or Novices because they hold greater predictive accuracy of their engagement with the material (Bruineberg et al., 2018; Friston, 2011, 2010, 2009; Friston et al., 2011, 2010, 2009), their motor control is likely more stable for task-crucial variables (Biryukova et al., 2015), and it is more difficult to significantly improve once a certain level of skill is reached (Gray and Lindstedt, 2017).

The effectiveness of this method for discriminating between bifaces was then put to a Blind Test of one Expert and one Intermediate knapper. Although, as I have noted above, the sample size of the Blind test is small, I would predict the following:

Hypothesis 4 (H^4):

- The quantitative methodology outlined will correctly distinguish between the two skill-levels in the Blind Test.

In addition to the analysis of the entire experimental assemblage, I performed a more specific analysis of the influence of shape on the effectiveness of accurate classification. I also scrutinised the effect of practice and instruction on one Novice (Experimental knapper 7), as well as the impact of instruction by an Expert on one Intermediate (Experimental knapper 4).

LDA analysis of the effects of biface planform shape (ovate or pointed) on classification accuracy was conducted because my experimental results were to be applied to an archaeological assemblage which exhibited a consistent ovate morphology. Essentially, this aimed to establish whether my criteria worked better for ovate or pointed handaxes in discriminating between known skill-levels. As a result, I would predict the following:

Hypothesis 5 (H^5):

- Handaxe planform shape (ovate or point) will affect the accuracy of the skill criteria for classifying handaxes correctly, because points tend to be less refined than ovates.

Separately, I conducted an LDA focused on how practice and instruction affected classification accuracy. This analysis consisted of two strands concentrating on two individuals – one Novice and one Intermediate. The first strand focused on myself as a beginner (Experimental knapper 7) making 33 bifaces over 6 months. I received tuition three times, producing 9 bifaces (27%). I also contributed bifaces made as part of a

knapping workshop where I was allowed to observe other practitioners of varying skill-levels (including Experts). One of the 5 bifaces produced during this session was completed with advice, but not direct instruction, from an Expert. The information imparted by the Expert was not specific regarding individual actions, but more general and hence qualified as advice and not step-by-step instruction. Therefore, 23 bifaces were made without instruction at different stages before, between, and after tuition. I investigated my development over time, but also the improvement generated via Expert instruction. Univariate statistics and LDA allowed me to assess my advancement in certain criteria, as well as how useful such gauges were in demonstrating the effect of tuition and/or practice. As introduced in Chapter 4, I would predict the following:

Hypothesis 6 (H^6):

- Among Novice knappers, practice will improve quantitative scores and evidence improvements in performance toward levels attained by Intermediates (Gray and Lindstedt, 2017). This will happen as the knapper learns to better control the delivery of the percussive blow.

Hypothesis 7 (H^7):

- Additionally, sudden significant improvements in skill-related criteria (e.g., Refinement ratios and Edge Straightness) will also be achieved through tuition by an Expert. This will occur via the introduction of league-stepping methods (Gray and Lindstedt, 2017).

The second strand concentrated on the Intermediate (Experimental knapper 4) who made handaxes over a long period and contributed 13 bifaces to the Intermediate sample, 9 bifaces to the Novice sample, and a further 25 which were deemed as 'Intermediate Drafts'. I contrasted the bifaces this participant made as a Novice with those made as an

Intermediate to see whether quantitative criteria were effective in discriminating between them. I also compared these bifaces to the handaxe they made under Expert instruction. Here, I would predict the following outcomes:

Hypothesis 8 (H^8):

- The bifaces that Experimental knapper 4 made in their third year of practice (as an Intermediate) will be quantitively distinguishable from those made in their first year of practice (as a Novice).

Hypothesis 9 (H^9):

- Quantitative analysis of bifaces made by Experimental knapper 4 under Expert step-by-step instruction will reveal a significant improvement in performance versus bifaces made without instruction.

6.5.1.2 Qualitative analyses

My qualitative analysis of the experimental assemblage was designed predominantly to give further detail, context, and explanations regarding why LDA misclassified some bifaces away from the skill-level category they belonged to. Where quantitative misclassifications occurred, detailed observations were provided regarding the non-quantifiable instances of high or poor skilled performance that a biface exhibited. Alternatively, comments concerning the effects that raw material quality or size may have had on quantitative scores are also included. This allowed context to be supplied regarding why a certain biface displayed quantitative indices more proportionate with a skill-level that did not belong to its maker.

A full qualitative breakdown for each biface in the experimental assemblage was not provided. Instead, I outline in the Results (Chapter 7) where the main misclassifications are and provide the qualitative context only in those few instances.

6.5.1.3 Questionnaire analyses

I compiled the answers given by all participants into Catalogue 2. This allowed for easier analysis of the responses. Quantitative data was assembled into tables, and responses to discussion-based questions were sorted by participant. Any pertinent comments and/or data have been woven into the theoretical framework (Chapter 4), Results (Chapter 7), and Discussion and Interpretation (Chapter 8).

6.5.2 Archaeological assemblage

The quantitative and qualitative analyses performed on the archaeological assemblage, as outlined at the start of this section, were similar to those implemented on the experimental assemblage. However, the prior information on the GZnivG bifaces was obviously different. Critically, the skill-level of biface makers from GZnivG can only be inferred. Therefore, the quantitative and qualitative analyses had to be used slightly differently to address the unknown aspect of the archaeological data and to provide a method for assessing skill-level differences at GZnivG by using the experimental data.

6.5.2.1 Quantitative analysis

Univariate statistical data was presented for the biface assemblage from GZnivG and compared with the experimental assemblage for each quantitative criterion. The

minimum and maximum scores, ranges, interquartile ranges, means, and standard deviations of all skill-related criteria were compared also. This conveyed a general comparative overview of GZnivG and the experimental pieces. It also allowed the differences in size and shape homogeneity previously discussed to be assayed.

LDA was used, but slightly differently to my experimental analysis because the skill-levels of the archaeological bifaces were unknown. Therefore, the most accurate experimental LDA results were used to examine the archaeological data. To interpret the GZnivG information, the data needed to be standardised to the distribution of all known skill levels, estimated from the experimental data. This was done by normalising experimental loadings and applying them to both experimental and archaeological data, which provided a method for quantitatively discriminating between the bifaces from GZnivG by skill-level thresholds discovered in the experimental data.

Normalisation of the experimental loadings allowed both experimental and archaeological datasets to be compared within the same frame of reference (i.e., the variance exhibited by different known skill-levels in the experimental assemblage). Using the experimental loadings without normalising them to assess the GZnivG assemblage, would have changed the archaeological data by introducing dimensional bias inherent in the experimental data. By removing the length of the experimental loading axes and keeping only the direction (i.e., variance) of them via normalisation, then the archaeological data could be assessed fairly by transposing it directly onto the frame of variance exhibited by the known experimental data.

A more detailed explanation of how LDAs and data normalisation were performed and presented can be found in Appendix 1, however, it is necessary here to summarise the process simply. The LDA performed on the experimental assemblage produced loading scores for each of the skill criteria. Loading scores represented the contribution each criterion made to the (direction of) each axis generated via the LDA. The loading scores for criteria can be either positive or negative but the largest loadings are the ones that dominate the axis direction (positive or negative). Each of the axes explains a percentage of the variance in the full data sample. An example is provided in Table 6.1 below.

	Axis 1 (81.99%)	**Axis 2 (14.57%)**	**Axis 3 (3.44%)**
Refinement	-2.4187	-9.7687	-0.65058
Th/L	6.4948	20.035	-9.8561
Profile Asymmetry	0.84593	4.7799	3.1278
Planform Asymmetry	8.7605	-5.0676	9.8187
Edge straightness	23.147	-5.0886	26.05

Table 6. 1. Example LDA Axis loading scores by criterion.

Biface ID	**Axis 1**	**Axis 2**	**Axis 3**
KL_1.1	-1.8424	-0.25096	0.036316
ADAMTURN.5	0.11844	0.61509	0.38562
LH_KL.1	0.40395	-1.2779	0.66086

Table 6. 2. Example LDA Axis score by biface.

These loading scores translate to ‘grades’ for each individual handaxe specimen on each axis (Table 6.2). For the experimental sample, the skill-level for each biface is determined. The LDA uses the scores for each handaxe to separate the different skill groups along the separate axes. With this, the experimental LDA was able to classify the parameters for Novice, Intermediate, and Expert ‘grades’ along each axis. These loading score parameters can be transferred and used on the archaeological assemblage. However, as mentioned above, the loading scores needed to be normalised to do so.

Normalisation involves dividing each loading in Table 6.2 by the length of the axis it belongs to. The length of each axis can be determined from Pythagoras’ theorem in higher dimensions, i.e., the length of axis 1 in Table 6.1 is equal to the square root of the sum of the squares of its loadings:

$$L_1 = \sqrt{(refinement_1)^2 + (Th/L_1)^2 + (ProfAsymm_1)^2 + (PlanAsymm_1)^2 + (EdgeStraight_1)^2}$$

Once the normalised experimental loadings were generated, they were used to transform the data for each criterion for both experimental and the archaeological handaxes. The measured values for each skill criterion for each handaxe was multiplied by their normalised loadings and then summed together to give the desired axis score for that specimen. This allowed normalised axis scores to be created for the known skill-level groups from the experimental assemblage and thresholds for Novice, Intermediate, and Expert to be designated on each axis. With this information it was possible to categorise the bifaces from GZnivG according to skill-level group.

However, as these thresholds are the result of human interpretation of the experimental data based on the range of scores visible on each axis for each skill-level category, it was necessary to conduct another LDA to check the preliminary classifications I had attributed to each individual GZnivG biface. This LDA assessed the accuracy of my classification based on normalisation and provided final quantitative classifications for each individual biface in the GZnivG assemblage. Accordingly, I would predict the following:

Hypothesis 10 (H^{10}):

- Two or more skill-levels will be discernible at GZnivG through quantitative analysis.

Moreover, as previously mentioned in Chapter 4, my prediction would be:

Hypothesis 11 (H^{11}):

- Because Novices and Experts constitute the extremes of the skill spectrum, they will be rarer than Intermediates.
    - My reasoning is based on the rationalisation that most handaxes at GZnivG will be made by those who are no longer Novices but have not accrued the relevant experience of practice to reach Expert levels of performance (i.e., >15 years). Moreover, as highlighted in Chapter 5, GZnivG bifaces were imported from outside the excavation area. If this relates to moving handaxes over significant distances, the reduced likelihood of Novices curating their poorly made bifaces would mean a signature of Novice handaxe manufacture is also less likely.

Finally, the univariate statistics of the bifaces by post-LDA classification grouping were presented to check whether there were certain criteria influencing classifications in a detrimental manner.

6.5.2.2 Qualitative analysis

The same qualitative criteria used to analyse the experimental assemblage was used for the GZnivG assemblage. Prior to quantitative data collection and analysis, qualitative observations and assessments of each biface were made. These allowed me to assign a preliminary assessment of skill-level (Novice, Intermediate, or Expert) to each biface. These observations were then incorporated into a catalogue (see Catalogue 1) alongside basic quantitative statistics. The catalogue outlines my qualitative interpretation of each individual bifaces' skill-level. Qualitative assessments for each biface were then correlated with their quantitative LDA classifications to understand how accurate the quantitative classifications were. I use this approach to predict the following conceivable outcomes:

Hypothesis 12 ($H^{12)}$:

- Qualitative assessment will, in the main, corroborate quantitative assessments.
- And, where misclassifications between the two occur, they will only occur between immediately adjacent skill levels (i.e., Novice and Intermediate, or Intermediate and Expert). There will be no qualitative assessment of Novice that is quantitatively classified as Expert or vice versa.

Hypothesis 13 (H^{13}):

- Comparison of qualitative assessment and quantitative classification will elucidate an elongation bias (McPherron, 1999) that will affect the accuracy with

which smaller, more heavily re-sharpened bifaces at Gouzeaucourt will be classified.

- Qualitative comparison will be key to demonstrating how significant the potential biasing effect of resharpening is.

6.5.2.3 Diacritical analysis

Lamotte's (2012, 2011) methodology required the assessment of the general sequence and direction of removals for each GZnivG biface using 8 pre-defined labels (see Figure 6.5). Each face of the handaxe was assessed and given a label. The final diacritical label was of the two faces added together (see Table 6.3). This amounted to 15 potential classifications.

Diacritical assessment
Unilateral and Unilateral
Unilateral and Bilateral
Unilateral and Trilateral
Unilateral and Centripetal
Unilateral and Diverse
Bilateral and Bilateral
Bilateral and Trilateral
Bilateral and Centripetal
Bilateral and Diverse
Trilateral and Trilateral

Trilateral and Centripetal
Trilateral and Diverse
Centripetal and Centripetal
Centripetal and Diverse
Diverse and Diverse

Table 6. 3. Diacritical assessment labels.

The assessment for each biface was then plotted against their skill-level classification. This was another source of data that could potentially differentiate knappers at GZnivG, and because GZnivG represents a shared repertoire of biface manufacture, I would predict the following:

Hypothesis 14 (H^{14}):

- Diacritical variation will not be significant.
- However, in line with my predictions in Chapter 4 regarding the less variable performance of more highly skilled knappers and Soriano's (2000) evidence for techno-functional consistency of flaking at GZnivG, Expert bifaces will yield fewer iterations of diacritical patterns than Intermediates and Novices.

6.6 Summary

This methodology outlines the objectives of my thesis clearly. I aim to use data from experimental investigation to analyse archaeological bifaces from GZnivG based on the skill-level of the knapper(s) who made them. The methods of analysis used were designed

to ascertain differences between bifaces at Gouzeaucourt that can be attributed to variations in skill-level. Given that the skill-level of GZnivG knappers is unknown, my mixed methods approach provides the necessary checks to ensure more secure interpretations.

7. Chapter 7 – Results

7.1 Introduction

The results of my experimental and archaeological investigations into assessing the skill-levels of Acheulean biface makers are presented below. The quantitative and qualitative criteria, methods, and analyses demonstrated here follow the Methodology outlined in Chapter 6. Here I submit my results with some cursory comments. A fuller discussion will follow in the next chapter.

7.2 Experimental Investigation

The experimental biface assemblage totalled 130 bifaces made by 7 individuals of varying ability. The breakdown of skill-levels and knappers that contributed to the assemblage is outlined in section 6.3.1 of my Methodology.

The handaxes were made with no restrictions (i.e., shape or size) placed upon their manufacture. This included the size of the initial raw material chosen for the task. The sole instruction given to each participant was that they were to make an Acheulean-style biface. The only potential biases to this occurred when both Novices and Intermediates were being occasionally tutored by more experienced knappers who may have preferred certain biface styles. However, this also allowed me to test how expert instruction impacted on the performance of less-skilled individuals.

7.2.1 Quantitative analysis

7.2.1.1 Univariate Results

I used Microsoft Excel, SPSS version 26.0, and PAST version 4.0 software to store and analyse the experimental quantitative data detailed in section 6.4.1 of my Methodology. Data tables for each of the criteria below is presented in Catalogue 4, Sheet 4.5.

1. Length (mm) and Breadth (mm):

Novices exhibited the greatest variations in biface length and, with Intermediate Drafts, were more likely to make smaller bifaces. Both Novices and Intermediate Drafts failed to make bifaces that were regularly above the mean of Intermediates or Experts (c.131mm).

The breadth range and standard deviation results showed that Experts and Intermediates normally were more consistent than Novices, producing a much tighter range with reduced variance. Again, as Questionnaire participant 3 suggests, Novices reduced too much of the biface edge.

The potential reasons why Novices did not, or could not, produce larger bifaces and often quite small bifaces is explained by Questionnaire participant 3 - learners regularly create comminution through incompletely propagating multiple fractures (see Catalogue 2, responses to Q16b).

2. Thickness (mm):

The ranges by skill-level showed that Novices and Intermediates produced thinner bifaces than the Expert's minimum score. Moreover, Intermediates had the lowest mean thickness. However, the ranges and standard deviations showed that Experts, again, were more consistent in their products (see Figure 7.1).

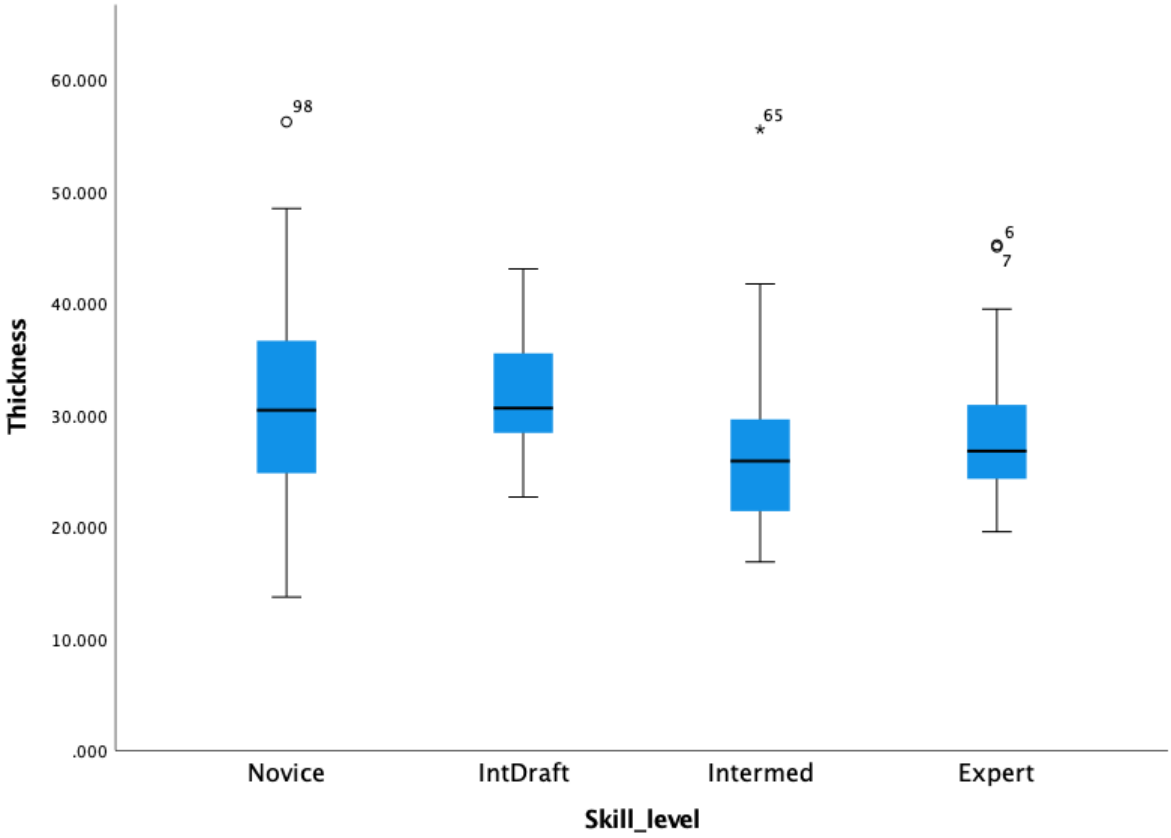


Figure 7. 1. Box and whisker plot showing range of experimental biface thickness by skill-level.

3. Weight (g):

The results were limited slightly in that the maximum weight possible on my scales was 500g. However, having an upper limit to weight did not obscure key differences by skill-level.

It should be re-iterated that the only instruction given was to make an Acheulean-style biface. The reasons for not imposing constraints on manufacture are outlined in Chapter 6 (section 6.3.1). As a result, occasionally Experts, and less frequently Intermediates, made bifaces on flakes. This has led to quite large ranges in weight.

Despite not making bifaces on flakes, Novices still behaved more variably than Experts and Intermediates. In addition, Intermediates acted more variably than Experts. This mirrored Length, Breadth, and Thickness statistics outlined above and when looked at together is potentially informative regarding the degree of control and flexibility that can be attributed to different skill-levels.

Novices utilising just one method of manufacture (i.e., reduction of a larger block) is probably a result of what they had been taught or witnessed more skilled individuals doing more regularly.

4. Breadth/Length ratio:

Breadth/Length ratio had previously been used to comment on the aesthetic capabilities and shape preferences of hominins (Pope et al., 2006). However, Gowlett (2011) showed it to be more directly related to biface length, finding that as length decreased, Breadth/Length ratio increased. Recent studies have supported Gowlett's (2011) findings (Hutchence and Debackere, 2018).

The Intermediate Draft mean was the highest of the skill-level categories. However, the Novice B/L ratio mean was more like Expert and Intermediate skill-levels, despite having

the lowest mean length and the lowest minimum length overall. In fact, counter to Gowlett's (2011) findings, the Novice B/L ratio mean was smaller than the Intermediate mean.

For the entire experimental assemblage there was a moderately strong statistically significant negative relationship between Length and B/L ratio: r_s (130) = -0.460, p<0.01. However, when broken down by skill-level, I found:

- Experts had a much stronger negative correlation: r_s (38) = -0.714, p<0.01
- Intermediate was stronger too at: r_s (20) = -0.621, p<0.01
- Intermediate Draft was slightly stronger at: r_s (25) = -0.526, p<0.01
- Novice displayed a weak negative correlation: r_s (47) = -0.300, p<0.05. Moreover, the correlation was significant at the 95% level not the 99% level as in Expert, Intermediate, and Intermediate Draft groups.

The reduced strength and significance of the correlation between Length and B/L ratio among Novices suggests they possessed less control over their actions regarding intentions related to dimensions and shape. This is a potential area for future study.

5. Roe's (1968) L1/L biface shape determination:

The main difference between Intermediate and Expert L1/L scores in comparison to Novice ones was that the narrower Intermediate range and greater Expert range was indicative of morphological choices made by the knappers, whereas greater range and standard deviation of Novice L1/L scores again suggested a lack of more precise control.

This was further borne out by the breakdown of shapes by skill-level:

- Novices produced 36% points, 58% ovates, and 6% cleavers. The latter were not cleavers in the traditional Acheulean sense of wide-tipped cleaving tools, but overly reduced, smaller bifaces with the max breadth nearer the tip.
- Intermediates displayed a 50/50 split between ovates and points. However, Figure 7.2 shows, the Intermediate range was very tight and 40% of Intermediate points were close to being categorised as ovates (within 5% of the threshold). If these were treated as attempted ovates, then the split changes to 70/30 in favour of ovates.
- Intermediate Drafts L1/L scores confirmed that Intermediate bifaces were attempted ovates. They revealed a 67/33 split in favour of ovates.
- The Expert split was 50/50 between ovates and points. However, this demonstrated the different preferences of individual Experts. Indeed, no metric points were within 5% of Roe's threshold for being categorised as ovate. Most notably, Experimental knapper 2 had a clear preference for making larger, heavier-butted pointed bifaces.

The narrow range of Intermediate L1/L scores could have resulted from a smaller sample size, however, Intermediate preference for making ovates is clear. No Intermediate attempted the large, pointed bifaces that for example Expert Experimental knapper 2 preferred to make.

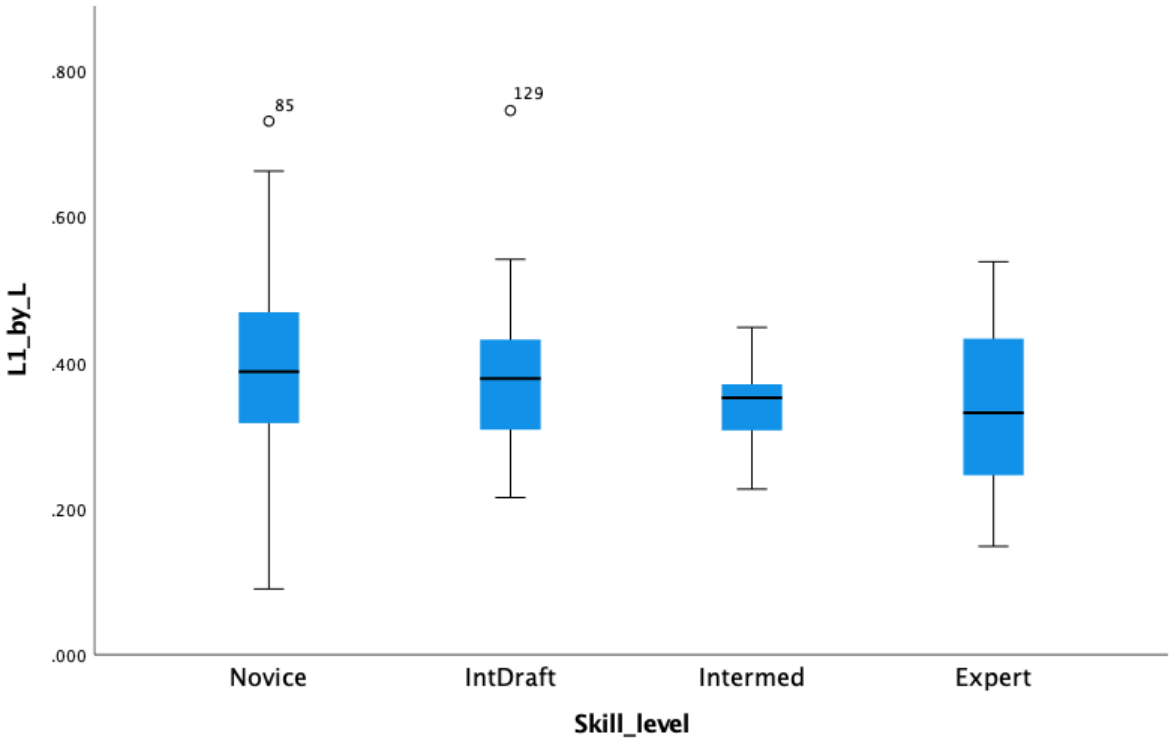


Figure 7. 2. Box and whisker plot showing range of experimental biface L1/L ratios by skill-level.

6. Ratio of Refinement:

The Ratio of Refinement is an index of how thick a biface is relative to its breadth (thickness/breadth). As outlined in my Methodology, it is a skill-related criterion because it quantitatively measures how well a biface has been thinned. It is measured along a scale of 0 to 1. Scores closer to zero indicate a more refined biface.

Intermediates and Experts created more refined bifaces more consistently. Their mean values were much lower, and their scores were less varied. Intermediate and Expert assemblages had high maximum refinement scores, but they were both outliers, with the next highest score being similar to the mean of the Novice assemblage.

The lowest minimum score and mean was provided by the Intermediate group. This otherwise surprising result may be indicative of the Intermediate sample being roughly half the size of the Expert one, although the lowest refinement index was still generated by the Intermediate. Moreover, the lowest Novice minimum score was below that of the Expert. Clearly, criteria exhibited overlaps between the skill-levels that, when taken in isolation, can affect discrimination between them.

Additionally, the less varied and slightly higher (than Intermediate) mean of the Experts may suggest that there was a preference among the Experts sampled for slightly thicker bifaces. This may have been for prehensile, ergonomic, or perceived functional butchery reasons, with overly thinned bifaces perhaps being more liable to breakage (Key and Lycett, 2017). Alternatively, it may reflect the greater variability in shapes produced by Experts, with more pointed bifaces being larger and thicker.

7. Thickness/Length ratio:

Like Ratio of Refinement, Thickness/Length ratio (Th/L) measures how well a knapper can thin a biface, whilst keeping the planview dimensions of the biface elongated, through correct edge maintenance.

Unsurprisingly, Th/L results were very similar to those of Ratio of Refinement. Expert and Intermediates could not be distinguished easily, whereas Novices had clear range, mean, and standard deviations separating them, including a lower minimum score

threshold which Novices did not reach. Notwithstanding this, there were intersections between the scores of all known skill-levels.

8. Planform and Profile Asymmetry:

Both criteria were measured along a scale of 0 to 1. Scores closer to zero reflect greater symmetry. The mean and standard deviation results for Planform Asymmetry implied that higher skill-levels (Intermediate and Expert) produced more symmetrical bifaces more consistently. Nevertheless, overlaps in the data were larger between all skill-levels for this criterion than the others discussed above.

Like Planform Asymmetry, the lower means, shorter ranges, and smaller standard deviations were made by higher skill-level groups (Intermediate and Expert) for Profile Asymmetry. Though, similar to Planform Asymmetry, there was significant overlap between all skill-levels. For example, the Novice minimum (i.e., most symmetrical) score was lower than the Intermediate minimum score and very close to the lowest score produced by Experts.

9. Edge Straightness:

Edge straightness measures how well a knapper maintains the straightness and sharpness of the biface edge for its intended cutting function (along a scale of 0 to 1). Scores closer to zero denote straighter biface edges.

This criterion was a good discriminator between skill-level groups. Minimum scores progressively improved, and ranges reduced, from lower to higher skill-levels. However, similarities in range between skill-levels obstructed it from being more confidently used on its own to discriminate between them.

10. Flake Scar number (>10mm):

Flake scar number was calculated by counting the total number of flake scars >10mm on both faces. Minimum and maximum scores, as well as means, showed that more flakes were attempted and/or removed as skill-level increased. However, Experts acted in a markedly more variable way than both Intermediates and Novices, removing a greater range of flakes. This likely relates to Experts making some bifaces on flakes and others via reducing whole nodules.

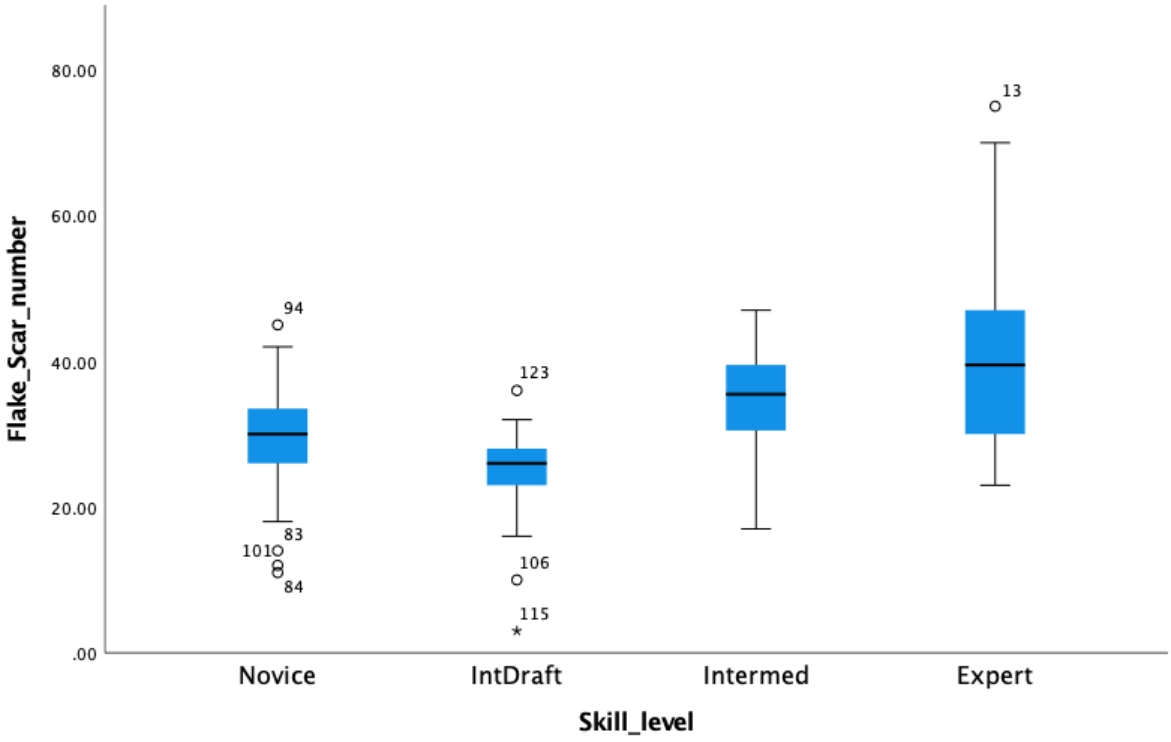


Figure 7. 3. Boxplot showing range of Flake Scar number scores by skill-level.

Overall, for all skill-related criteria (Ratio of refinement, Thickness/Length ratio, Planform and Profile Asymmetry, Edge straightness, and Flake Scar number) presented above, the means of both Intermediate and Expert groups did not differ by more than both groups' standard deviation/variance. Whereas, except for Profile Asymmetry, this was generally the case between Novice and Expert or Intermediate mean scores. These results suggest that differentiating between Intermediate and Expert bifaces in subsequent analyses will be more difficult than clearly distinguishing between them and Novice bifaces.

7.2.1.2 Bivariate correlations

The purpose of this section was to demonstrate whether skill-related criteria correlated in a statistically significant way. For all skill-criteria, more competent scores were closer to zero. Consequently, a key question is: did more/less skilful scores in one criterion correlate significantly with more/less proficient scores in another criterion?

Spearman's Rank r_s correlation coefficients were undertaken in IBM SPSS Statistics (version 26.0) between the skill-related criteria for the entire experimental biface assemblage, and for individual skill-levels. Scatter plots for correlations are presented in Catalogue 4, Sheet 4.6.

The rationale behind using Spearman's Rank, and not Pearson's, is outlined in section 6.5.1.1 in my Methodology. Statistical significance was designated at both 95% and 99% confidence levels, in line with generally accepted significance thresholds. Strong correlations were indicated where $r_S > 0.5$.

Each correlation is presented in two parts. One will include Intermediate Drafts and the other will exclude them. This was to compare the impact of drafted bifaces, that were deemed inadequate, or unfinished, by the Intermediate, on correlation strength and significance.

7.2.1.2.1 Results

1. Ratio of Refinement and Thickness/Length

Sample including Intermediate Drafts

- Strong, positive correlation: r_S (130) = 0.780, p <0.01.

Sample excluding Intermediate Drafts

- Strong, positive correlation: r_S (105) = 0.838, p <0.01.

Both correlations corroborated Hutchence and Debackere's (2018) findings for archaeological handaxe assemblages in Britain: (r_S (164) = 0.803, p <0.05.

The scatter graph (see Catalogue 4, Sheet 4.6) shows that an overlap existed between the skill-levels. However, both Expert and Intermediate handaxes were generally closer to zero on either axis than both Novice and Intermediate Draft bifaces. When split according to skill-level category, however, the correlations were:

- Expert: r_S (38) = 0.684, p <0.01.
- Intermediate: r_S (20) = 0.951, p <0.01.
- Intermediate Draft: r_S (25) = 0.209, p = 0.316.
- Novice: r_S (47) = 0.768, p <0.01.

Strong correlations between two criteria measuring the knapper's ability to thin bifaces was expected for all skill-levels. Therefore, it seems curious that Intermediate Draft bifaces do not show a statistically significant correlation between the Refinement ratios. This result likely reflects the incomplete, drafted nature of these bifaces.

2. Ratio of Refinement v Edge straightness

Sample including Intermediate Drafts

- Strong, positive correlation: r_S (130) = 0.502, p <0.01.

Sample excluding Intermediate Drafts

- Strong, positive correlation: r_S (105) = 0.513, p <0.01.

Both supported Hutchence and Debackere's (2018) results: r_S (148) = 0.501, p <0.05.

The scatter graph (again, see Catalogue 4, Sheet 4.6 for all scatter graphs) shows the main division between skill groups was on the Y axis in relation to Edge Straightness. When sub-divided by skill-level, however, the correlations were:

- Expert: r_S (38) = -0.023, p = 0.889.
- Intermediate: r_S (20) = 0.623, p <0.01.
- Intermediate Draft: r_S (25) = 0.217, p = 0.298.
- Novice: r_S (47) = 0.518, p <0.01.

The very weak and highly insignificant correlation between these two criteria for Expert bifaces suggests that Experts could regularly produce straight edges regardless of the extent to which they refined their bifaces. However, comparison of Expert bifaces with Intermediate ones shows that the Edge straightness means of both groups did not differ by more than both groups' variance (standard deviation). Thus, both groups exhibited similar corresponding scores that were virtually indistinguishable. Therefore, it is intriguing that Intermediate bifaces had a strong and significant relationship between these two variables, despite Edge straightness being consistent among this group.

The result for Experts is likely driven by the greater (versus Intermediate bifaces) variance in biface shape as highlighted in the L1/L statistics (see Catalogue 4, Sheet 4.5, Table 6). Certainly, pointed bifaces made by Experts were largely thicker and less refined but still exhibited very straight edges, whereas Intermediates did not make large, thick pointed bifaces. Therefore, because Intermediates predominantly produced ovates, when they attained a poor Refinement score it correlated more with poor Edge straightness

scores because it reflected poor performance rather than a deliberate design choice by the knapper. The same applies for Novice bifaces. However, the insignificant (but not as insignificant as the Expert) result for Intermediate Draft bifaces again likely conveys the particular stages of manufacture at which the Intermediate abandoned fabrication.

3. Thickness/Length ratio vs Edge straightness

Sample including Intermediate Drafts

- Strong, positive correlation: r_S (130) = 0.599, p <0.01.

Sample excluding Intermediate Drafts

- Strong, positive correlation: r_S (105) = 0.577, p <0.01.

Both upheld Hutchence and Debackere's (2018) results: r_S (147) = 0.6, p <0.05.

The scatter graph shows the main division between skill groups on the Y axis in relation to Edge Straightness. When separated by skill-level, however, the correlations were:

- Expert: r_S (38) = 0.053, p = 0.754.
- Intermediate: r_S (20) = 0.69, p <0.01.
- Intermediate Draft: r_S (25) = 0.456, p <0.05.
- Novice: r_S (47) = 0.59, p <0.01.

The significance of this correlative breakdown by skill-level mirrors the comments above for the correlation between Ratio of refinement and Edge straightness.

4. Ratio of Refinement and Planform Asymmetry

Sample including Intermediate Drafts

- Moderate, positive correlation: r_S (130) = 0.460, p <0.01.

Sample excluding Intermediate Drafts

- Moderate, positive correlation: r_S (105) = 0.486, p <0.01.
    - Very near to the threshold for strong correlations of 0.5.

Both were close to verifying Hutchence and Debackere's (2018) results: r_S (141) = 0.448, p <0.05. However, corroboration of their results was closer when Intermediate Drafts were included. When organised by skill-level, however, the correlations were:

- Expert: r_S (38) = 0.547, p <0.01.
- Intermediate: r_S (20) = 0.398, p = 0.082.
- Intermediate Draft: r_S (25) = 0.232, p = 0.264.
- Novice: r_S (47) = 0.248, 0.093.

The breakdown by skill-level suggests that Expert bifaces drove the significance of the overall assemblage correlation. Again, this was probably due to the increased variance in morphology among Experts (compared to Intermediates). The thicker, larger, and less refined pointed Expert bifaces had more asymmetrical butts than the ovates made by other Experts and Intermediates. Therefore, this significant correlation among Experts reflects intentional handaxe shape choices.

Novices possessed a greater range of both Refinement and Planform Asymmetry than Experts but did not show a significant correlation between the two criteria. This indicates that the shape they formed was less purposeful and unrelated to the control the knappers had over refinement, symmetry, and ultimate morphology.

Intermediate ovates, on the other hand, did not express a correlation between these two criteria because of the stable and less varied handaxe forms they made. This intimates that when biface shape is consistent, Planform Asymmetry has no relation to the degree of refinement a knapper can produce.

5. Thickness/length ratio vs Planform Asymmetry

Sample including Intermediate Drafts

- Low to moderate, positive correlation: r_S (130) = 0.382, p <0.01.

Sample excluding Intermediate Drafts

- Low to moderate, positive correlation: r_S (105) = 0.426, p <0.01.

Correlations between these two variables more clearly corroborate Hutchence and Debackere's (2018) results [r_S (140) = 0.365, p <0.05] when Intermediate Drafts are included. When separated by skill-level, however, the correlations were:

- Expert: r_S (38) = 0.413, p <0.05.
- Intermediate: r_S (20) = 0.312, p = 0.180.
- Intermediate Draft: r_S (25) = -0.033, p = 0.877.

- Novice: r_S (47) = 0.175, p = 0.240.

The significance of this correlative breakdown by skill-level mirrors the comments above for the correlation between Ratio of refinement and Planform Asymmetry.

6. Ratio of Refinement vs Profile Asymmetry

Sample including Intermediate Drafts

- Weak, positive correlation: r_S (130) = 0.181, p <0.05.

The statistical significance (95%) and strength of this correlation was weaker than other correlations presented so far.

Sample excluding Intermediate Drafts

- No correlation: r_S (105) = 0.365, p =0.105.

When organised by skill-level, however, the correlations were:

- Expert: r_S (38) = 0.266, p = 0.106.
- Intermediate: r_S (20) = -0.122, p = 0.609.
- Intermediate Draft: r_S (25) = 0.228, p = 0.274.
- Novice: r_S (47) = 0.120, p = 0.423.

This breakdown by skill-level demonstrates that no skill-level group had a significant correlation between these two criteria. This upholds the overall assemblage results that show these two criteria do not correlate in any meaningful way.

7. Thickness/Length ratio vs Profile Asymmetry

Sample including Intermediate Drafts

- No correlation: r_S (130) = 0.098, p =0.267.

Sample excluding Intermediate Drafts

- No correlation: r_S (105) = 0.09, p =0.362.

The lack of correlations between both refinement ratios and Profile Asymmetry outlined above insinuates that Profile Asymmetry was not related to how well a biface was thinned. Well thinned and poorly thinned bifaces equally exhibited high- or low-Profile Asymmetry scores.

I did not expect this, given that Profile symmetry requires greater understanding of edge maintenance, managing the plane of intersection between the two faces separated by the cutting edge, as well as creating equal convexities of both faces in profile relative to the cutting edge. Hence, good Profile Asymmetry is a much more difficult outcome to achieve than planform symmetry because more variables need to be controlled by the knapper. As Questionnaire participant 4 (see Catalogue 2, Q16b) observed, they 'struggled to control the secant plane' as a learner. When sub-divided by skill-level, however, the correlations were:

- Expert: r_S (38) = 0.368, p <0.05.
- Intermediate: r_S (20) = -0.221, p = 0.349.

- Intermediate Draft: r_S (25) = -0.173, p = 0.409.
- Novice: r_S (47) = -0.075, p = 0.618.

The significance of this correlative breakdown by skill-level largely mirrors the comments above for the correlation between Ratio of refinement and Planform Asymmetry. The main difference was that Expert bifaces possessed a low to moderate, statistically significant, correlation between the criteria. This again is probably due to the increased variance in deliberate handaxe morphology among Expert bifaces. Thicker, larger pointed bifaces with higher Th/L ratio scores may have usually produced more asymmetrical profiles than ovates, however, the correlation is not strong.

8. Planform Asymmetry vs Edge straightness

Sample including Intermediate Drafts

Moderate, positive correlation: r_S (130) = 0.432, p <0.01.

Sample excluding Intermediate Drafts

- Moderate, positive correlation: r_S (105) = 0.443, p <0.01.

There was a stronger correlation between these variables than in Hutchence and Debackere's (2018) results: r_S (141) = 0.352, p <0.05.

The scatter plot shows that, in general, Experts and Intermediates had straighter edges and more symmetrical bifaces. When separated by skill-level, however, the correlations were:

- Expert: r_S (38) = -0.082, p = 0.623.
- Intermediate: r_S (20) = 0.432, p = 0.057.
- Intermediate Draft: r_S (25) = 0.138, p = 0.509.
- Novice: r_S (47) = 0.433, p <0.01.

The highly insignificant correlation between these criteria among Expert bifaces was perhaps indicative of the reduced importance of Planform Asymmetry for Experts. Experts worked within a tight range of Edge straightness but were very variable symmetrically. The same applied to Intermediate bifaces, although their correlation was close to the 95% confidence threshold.

Novice bifaces, on the other hand, produced a moderate, statistically significant correlation. This was perhaps in part due to greater variance in ability between those knappers categorised as Novices than among knappers of other skill-levels (see Tables 9 and 11, Sheet 4.5, Catalogue 4). For instance, Experimental knapper 4's (Novice) mean Edge Straightness score was 0.07 and (Novice) Experimental knapper 6's was 0.11. Similarly, Experimental knapper 4's mean Planform Asymmetry score was 0.06, while Experimental knapper 6's was 0.11. Whereas (Expert) Experimental knapper 1's mean Edge Straightness score was 0.055 and (Expert) Experimental knapper 2's was 0.051. In addition, their Planform Asymmetry mean scores were similar too, at 0.085 (Experimental knapper 1) and 0.10 (Experimental knapper 2).

9. Profile Asymmetry vs Edge Straightness

Sample including Intermediate Drafts

No correlation: r_S (130) = 0.149, p = 0.09.

Sample excluding Intermediate Drafts

- No correlation: r_S (105) = 0.137, p = 0.163.

Planform Asymmetry correlating with Refinement and Edge straightness is predictable, but Profile Asymmetry not doing so, was an unexpected result. However, the scatter plot is useful for visualising the difference in Edge straightness scores between the skill-levels. When organised by skill-level, the correlations were:

- Expert: r_S (38) = 0.124, p = 0.458.
- Intermediate: r_S (20) = -0.301, p = 0.198.
- Intermediate Draft: r_S (25) = 0.048, p = 0.821.
- Novice: r_S (47) = -0.015, p = 0.923.

This breakdown by skill-level demonstrates that no skill-level group had a significant correlation between these two criteria. This supports the overall assemblage results that these variables do not correlate in any meaningful way.

<u>10. Flake Scar number (>10mm) vs Ratio of Refinement</u>

Sample including Intermediate Drafts

- Weak, negative correlation: r_S (130) = -0.307, p <0.01.

Sample excluding Intermediate Drafts

- Weak, negative correlation: r_S (105) = -0.298, p <0.01.

However, when sub-divided by skill-level, the correlations were:

- Expert: r_S (38) = -0.412, p <0.05.
- Intermediate: r_S (20) = 0.191, p = 0.419.
- Intermediate Draft: r_S (25) = -0.248, p = 0.231.
- Novice: r_S (47) = -0.105, p = 0.484.

The breakdown by skill-level suggests that Expert bifaces drove the significance of the weak overall assemblage correlation. Experts removed more flakes more often than other skill-levels, and Expert bifaces that revealed the most flake removals were often among the most refined bifaces. Yet, the variance of Expert Flake Scar number data is far greater (see Catalogue 4, Sheet 4.5, Table 12) than for other skill-levels groups. Again, I propose that the greater variance in biface morphology among Expert bifaces had an impact on the amount of flakes the Experts removed.

11. Flake Scar number (>10mm) vs Thickness/Length ratio

Sample including Intermediate Drafts

- Weak to moderate, negative correlation: r_S (130) = -0.385, p <0.01.

Sample excluding Intermediate Drafts

- Weak, negative correlation: r_S (105) = -0.344, p <0.01.

Correlations between both refinement ratios and Flake Scar number broadly suggests that Experts and Intermediates created more refined handaxes that displayed a greater number of removals. However, the scatter plot shows overlap between the skill-levels. When separated by skill-level, however, the correlations were:

- Expert: r_S (38) = -0.454, p <0.01.
- Intermediate: r_S (20) = 0.159, p = 0.503.
- Intermediate Draft: r_S (25) = -0.167, p = 0.425.
- Novice: r_S (47) = -0.185, p = 0.214.

The significance of this correlative breakdown by skill-level reinforces the comments above regarding the correlation between Flake Scar number and Ratio of refinement.

12. Flake Scar number vs Edge straightness

Sample including Intermediate Drafts

- Strong, negative correlation*:* r_S (130) = -0.527, p <0.01.

Sample excluding Intermediate Drafts

- Strong, negative correlation*:* r_S (105) = -0.499, p <0.01.

This correlation generally implies that Expert and Intermediate handaxes presented straighter edges and more removals. When organised by skill-level, however, the correlations were:

- Expert: r_S (38) = -0.375, p <0.05.
- Intermediate: r_S (20) = 0.200, p = 0.397.
- Intermediate Draft: r_S (25) = -0.670, p <0.01.
- Novice: r_S (47) = -0.463, p <0.01.

The significant negative correlation between these variables for Expert bifaces suggests that handaxes displaying fewer removals had more sinuous edges. However, this is not a strong correlation and Edge straightness seems to be largely stable for Experts regardless of Flake Scar number.

For Intermediate bifaces there was no correlation, suggesting both Edge straightness and Flake Scar number data was similar for all Intermediate bifaces. However, Intermediate Draft bifaces showed a strong and significant negative correlation. This may indicate how soon into manufacture the Intermediate abandoned production. Though, this cannot be confirmed by my experimental results. Nonetheless, the moderate to strong, significant negative correlation among Novice bifaces is reflective of the greater variance in Edge Straightness scores among Novice knappers, as outlined in correlation 8 above (and in Table 11, Sheet 4.5, Catalogue 4).

13. Flake Scar number vs Planform Asymmetry

Sample including Intermediate Drafts

- Moderate, negative correlation: r_S (130) = -0.401, p <0.01.

Sample excluding Intermediate Drafts

- Moderate, negative correlation: r_S (105) = -0.413, p <0.01.

This correlation indicates that Experts and Intermediates broadly created more symmetrical bifaces with more removals than did Novices. However, Flake Scar number was the more dominant discriminating variable in these cases. When separated by skill-level, however, the correlations were:

- Expert: r_S (38) = -0.352, p <0.05.
- Intermediate: r_S (20) = 0.003, p = 0.990.
- Intermediate Draft: r_S (25) = -0.282, p = 0.172.
- Novice: r_S (47) = -0.328, p <0.05.

The moderately strong and significant negative correlation between these variables for Expert bifaces suggests that those displaying more removals were more symmetrical. However, the significance of this correlation is driven by two of the most symmetrical bifaces which also exhibited the most flake removals. Without these two handaxes (bifaces DK1 and KL_1.1), the Expert distribution would exhibit lower Flake Scar number variance and resemble more that of the Intermediate category (see Table 12, Sheet 4.5, Catalogue 4), which showed no correlation.

Again, similar to correlation number 8, the significance of the Novice correlation may be attributed to the greater variance in Planform Asymmetry scores exhibited by Novices than other skill-levels (see Table 9, Sheet 4.5, Catalogue 4). Although, like the Expert correlation, the Novice correlation is weak due to variance in Flake Scar number for this category being lower.

14. Flake Scar number vs Profile Asymmetry

Sample including Intermediate Drafts

- Weak, negative correlation: r_S (130) = -0.312, p <0.01.

Sample excluding Intermediate Drafts

- Weak, negative correlation: r_S (105) = -0.251, p <0.01.

The statistical significance and strength of this correlation was influenced by the more dominant discriminating variable of Flake Scar number. When organised by skill-level, the correlations were:

- Expert: r_S (38) = -0.379, p <0.05.
- Intermediate: r_S (20) = -0.302, p = 0.195.
- Intermediate Draft: r_S (25) = -0.291, p = 0.158.
- Novice: r_S (47) = -0.114, p = 0.445.

Expert bifaces clearly impacted the significance of the overall assemblage's weak correlation. Again, it seems probable that the two outlying Expert bifaces with the highest number of flake scars affected the results. The other skill-levels show that these two variables do not correlate strongly or significantly.

7.2.1.2.2 Summary

Univariate and bivariate analyses highlighted that Profile Asymmetry was not a good indicator for determining skill-level, whereas Edge Straightness and Refinement ratios were superior in discriminating between skill-level categories. In general, however, each criterion and correlation showed that Novices could be separated from Intermediate and Expert bifaces, despite there being some overlap between all skill-levels.

Separating Intermediate and Expert bifaces was much more difficult, however. Therefore, I conducted a multivariate Linear Discriminant Analysis (LDA), a form of Discriminant Function Analysis (DFA). Outlined in the following section, this went beyond the previous analyses by examining the relationship between the skill-related criteria according to known skill-level group. LDA identifies occurrences of scores by skill-level category and predicts whether the scores a biface achieves are 'truthful' to the group it has been assigned.

Furthermore, because handaxe shape seemed to affect the correlations of skill-related criteria, I also conducted LDA via biface shape to assess the effect it had on the classificatory accuracy of my criteria. This was necessary because GZnivG bifaces, as already discussed, are predominantly ovate. Thus, I wanted to establish the best criteria for correctly classifying bifaces of this morphological type.

Separately, I conducted a LDA of Experimental knapper 7 to establish whether the overall Novice results have been influenced by instruction, in addition to the marked differences in ability between the Novice knappers in my investigation (as highlighted above). I also conducted LDA to assess the effect of practice on Intermediate quantitative scores.

7.2.1.3 Multivariate analysis

The process and rationale behind normalisation are outlined in Chapter 6, section 6.5.2.1. Fuller mathematical detail of the method for conducting and reporting this LDA is provided in Appendix 1. Like the bivariate correlations above, each LDA is presented in two parts. One includes Intermediate Drafts and the other excludes them. Again, this was to evaluate the impact that draft bifaces had on correct skill-level classification. To guard against overestimating the predictive power of the LDA, jackknifing was performed. Jackknifing omitted one individual biface in the sample and generated an axis using the remaining data (i.e., minus the biface removed) that was then used to classify the original biface. This meant the biface's data was not being used in its own classification.

7.2.1.3.1 Results

The skill criteria that I have used before (cf. Hutchence and Debackere, 2018) included both Refinement ratios, Planform Asymmetry, and Edge straightness. They have all correlated statistically significantly with each other and have broadly shown, unsurprisingly, that more highly skilled individuals achieve better scores. Therefore, they were used as the main four LDA criteria here. In addition, Profile Asymmetry and Flake scar number were tested separately and together in conjunction with the main four criteria to ascertain their impact on LDA skill-classification accuracy.

As I outlined in my Methodology, however, a key assumption inherent in a Linear Discriminant Analysis is that each measured variable, or skill-related criterion (e.g., Ratio

of refinement), has identical covariance within discrete groups (e.g., skill-level category). However, if the different groups' distribution in each criterion/variable does not have the same covariance, a separation of the groups can still be performed if the measured values for the independent groups differ enough (i.e., if the means of the individual groups diverge by more than both groups' standard deviation/variance).

In Catalogue 4, I have shown that covariance between skill groups for each criterion is not identical. Yet, although the means of the separate groups for each variable generally deviated by more than each groups' variance between higher and lower skill-levels, this was not the case between Intermediate and Expert (see Catalogue 4, Sheet 4.5, Tables 7 – 12). This resulted, as I demonstrate below, in aspects of my Hypotheses 1 – 3 not being able to be substantiated via LDA. Nevertheless, this method was still effective in separating Novices from both Expert and Intermediate bifaces, and vice versa.

LDA 1: Ratio of refinement, Thickness/Length ratio, Edge straightness, and Planform Asymmetry

Sample including Intermediate Drafts (n = 130)

- 50% jackknifed classification accuracy.

Loadings

In Axis 1, which explained most of the variance, the strongest discriminator was Edge Straightness and Refinement was the weakest discriminator. However, across three axes, Planform Asymmetry was weakest.

	Axis 1 (80.73%)	Axis 2 (15.38%)	Axis 3 (3.89%)
Refinement	-1.8381	-9.8427	-8.5483
Edge Straightness	21.012	-11.932	26.961
Plan Asymmetry	4.9367	-1.405	-4.2834
Th/L	7.6287	22.94	-1.7732

Table 7. 1. LDA 1: Skill criteria loadings (including Intermediate Drafts).

Confusion Matrix

The below table shows the number of bifaces by skill-level group which have been both correctly assigned to their skill group and incorrectly assigned to other skill-level categories. The main areas of overlap were between skill-level groups directly adjacent to each other. Experts were predominantly misclassified with Intermediates; in fact, more of the Expert bifaces were re-classified as Intermediate than Expert. However, Intermediate misclassification was evenly split between Experts and lower skill-levels. Intermediate Drafts, on the other hand, were mostly misclassified with Novices, and Novices were mainly misclassified as Intermediate Drafts and Intermediates but not with Experts.

	Novice	Intermediate Draft	Intermediate	Expert	Total
Novice (n=47)	18	13	11	5	47
Intermediate Draft (n=25)	4	17	2	2	25

Intermediate (n=20)	1	2	14	3	20
Expert (n=38)	1	4	17	16	38
Total (n=130)	24	36	44	26	130

Table 7. 2. LDA 1: Confusion Matrix (including Intermediate Drafts).

Scatter plot

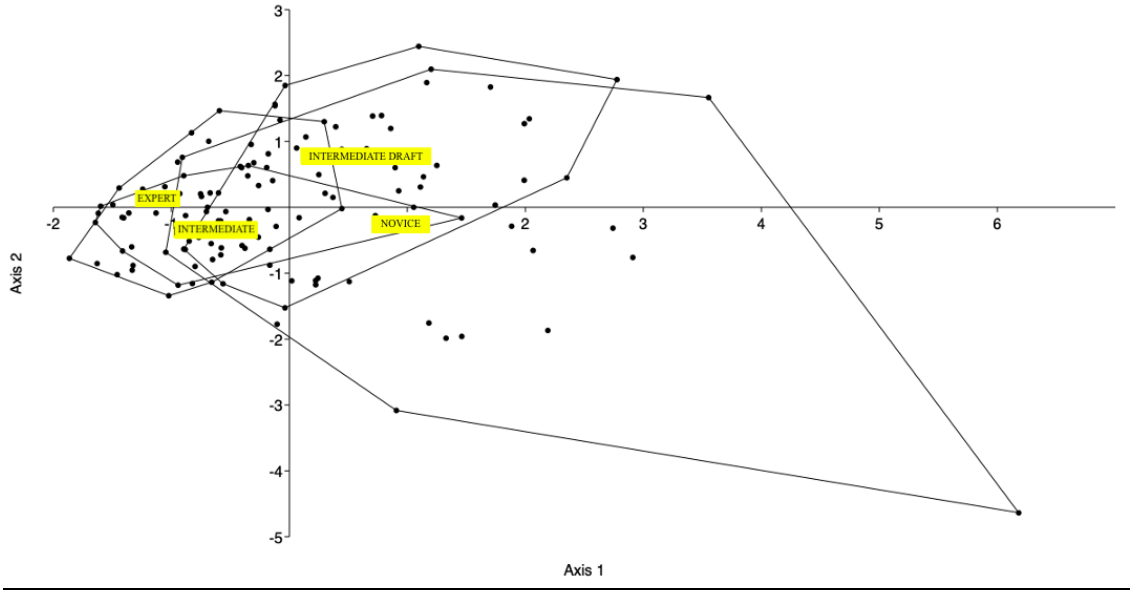


Figure 7. 4. Scatter plot of LDA 1 jackknifed experimental biface scores grouped by reclassified skill-level (including Intermediate Drafts).

Sample excluding Intermediate Drafts (n = 105)

- 58.1 jackknifed classification accuracy.

Loadings

The strongest discriminator again was Edge straightness, and the weakest discriminative criterion was Refinement.

	Axis 1 (95.39%)	Axis 2 (4.61%)
Refinement	-2.0419	-6.5247
Edge Straightness	21.45	26.212
Planform Asymmetry	5.1699	-4.1772
Th/L	6.7935	-4.59

Table 7. 3. LDA 2: Skill criteria loadings (excluding Intermediate Drafts).

Confusion Matrix

When Intermediate Drafts were removed from the sample, LDA accuracy improved by over 8 percentage points. The main improvements were the number of Novices an Experts correctly assigned. However, the removal of Intermediate Drafts did not affect the number of Intermediate bifaces correctly classified.

Clearly, when Intermediate Drafts were included (above) they mostly interrupted exact classification for Experts and Novices. When Intermediate Drafts were excluded, the 4 Expert and 2 Intermediate bifaces misclassified as Intermediate Drafts were reclassified as Expert and Novice. More significantly, the 13 Novice bifaces misclassified as Intermediate Drafts were reclassified as mostly Novice (11) and 2 Intermediate.

Skill-level	Novice	Intermediate	Expert	Total
Novice (n=47)	29	13	5	47

Intermediate (n=20)	2	14	4	20
Expert (n=38)	3	17	18	38
Total (n=105)	34	44	27	105

Table 7. 4. LDA 1: Confusion Matrix (excluding Intermediate Drafts).

Scatter Plot

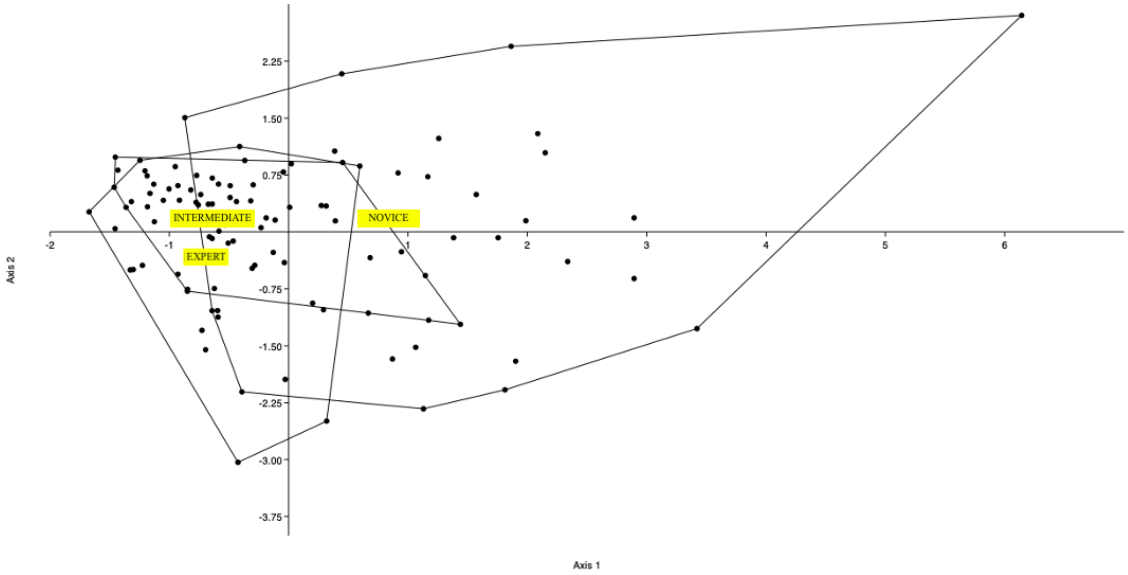


Figure 7. 5. Scatter plot of LDA 1 jackknifed experimental biface scores grouped by reclassified skill-level (excluding Intermediate Drafts).

Crucially, as envisaged at the start of this section, both LDAs (including and excluding Intermediate Drafts) showed that although Expert and Intermediate skill-levels were more difficult to discriminate between, both were more easily differentiated versus Novice. This is clearer on the scatter plots (see Figures 7.4 and 7.5).

LDA 2: Adding Profile Asymmetry

Sample including Intermediate Drafts (n = 130)

- 47.69% jackknifed classification accuracy.

Loadings

Profile Asymmetry was the weakest discriminator in Axis 1 and 3 and was a close second weakest to Planform Asymmetry in Axis 2.

	Axis 1 (79.71%)	**Axis 2 (16.78%)**	**Axis 3 (3.51%)**
Refinement	-3.0932	-9.6304	-7.3463
Edge Straightness	20.203	-12.862	26.898
Planform Asymmetry	3.8244	-3.171	-4.6405
Th/L	9.5204	20.314	-4.5575
Profile Asymmetry	2.9612	3.3974	1.3899

Table 7. 5. LDA 2: Skill criteria loadings (including Intermediate Drafts).

Confusion Matrix

Predictably, given the univariate and bivariate results above, the addition of Profile Asymmetry increased misclassification versus LDA1. Two Expert handaxes, originally classified as Intermediate in LDA1, were reclassified as Intermediate Draft – which for the main four skill criteria was mostly closer to Novice than Intermediate scores (see Catalogue 4, Sheet 4.5). Also, three bifaces originally correctly classified as Intermediate in LDA1 were reclassified as Expert (2) and Novice (1) in LDA2.

	Novice	Intermediate Draft	Intermediate	Expert	Total
Novice (n=47)	18	13	11	5	47
Intermediate Draft (n=25)	5	17	1	2	25
Intermediate (n=20)	2	2	11	5	20
Expert (n=38)	1	6	15	16	38
Total (n=130)	26	38	38	28	130

Table 7. 6. LDA 2: Confusion Matrix (including Intermediate Drafts).

Scatter plot

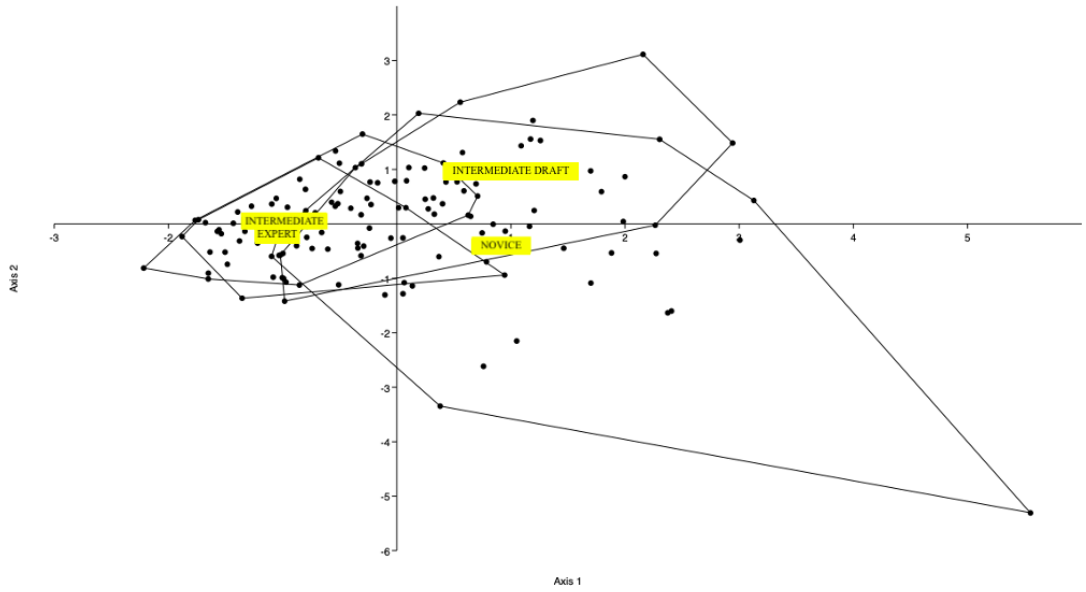


Figure 7. 6. Scatter plot of LDA 2 jackknifed experimental biface scores grouped by reclassified skill-level (including Intermediate Drafts).

Sample excluding Intermediate Drafts (n = 105)

- 53.33% jackknifed classification accuracy.

Loadings

Profile Asymmetry was again the weakest discriminator, followed closely by Refinement.

	Axis 1 (95.74%)	**Axis 2 (4.26%)**
Refinement	-2.7573	-6.7082
Edge Straightness	21.357	25.095
Planform Asymmetry	4.2633	-4.7131
Th/L	7.7375	-4.268
Profile Asymmetry	2.7516	1.3406

Table 7. 7. LDA 2: Skill criteria loadings (excluding Intermediate Drafts).

Confusion Matrix

In comparison to when Intermediate Drafts were excluded for LDA 1, the classification of Intermediates was more strongly affected with the addition of Profile Asymmetry when Intermediate Drafts were excluded for LDA 2:

- 6 Expert bifaces misclassified as Intermediate Drafts were reclassified as Novice (3), Expert (2), and Intermediate (1).
- 2 Intermediate bifaces misclassified as Intermediate Drafts were reclassified as Intermediate and Novice.
- 13 Novice bifaces misclassified as Intermediate Drafts were reclassified as Novice (8), Intermediate (4), and Expert (1).

	Novice	Intermediate	Expert	Total
Novice (n=47)	26	15	6	47
Intermediate (n=20)	3	12	5	20
Expert (n=38)	4	16	18	38
Total (n=105)	33	43	29	105

Table 7. 8. LDA 2: Confusion Matrix (excluding Intermediate Drafts).

Scatter plot

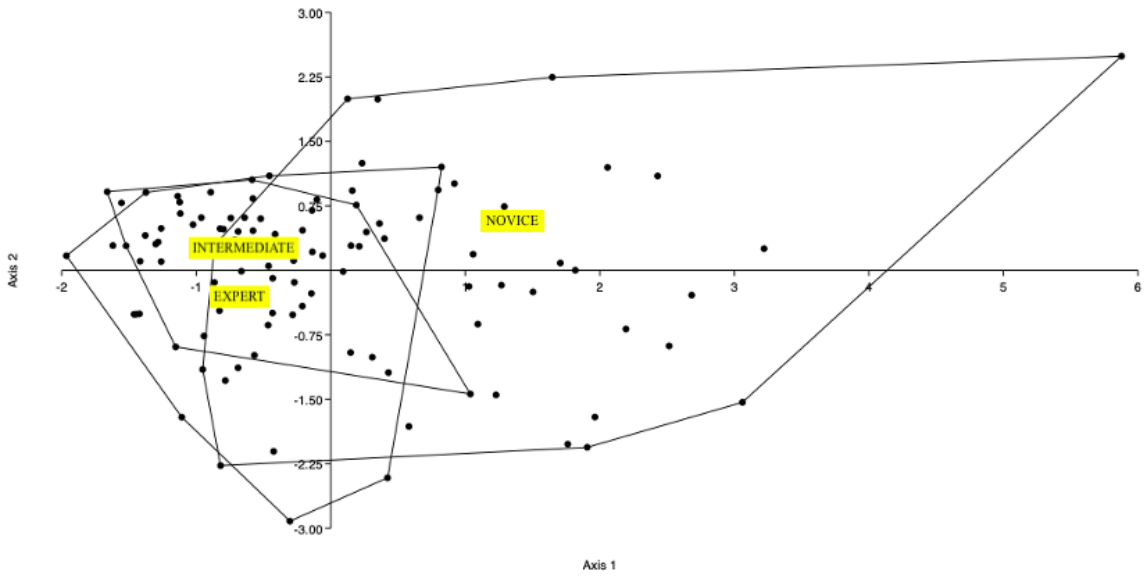


Figure 7. 7. Scatter plot of LDA 2 jackknifed experimental biface scores grouped by reclassified skill-level (excluding Intermediate Drafts).

Overall, the addition of Profile Asymmetry negatively affected correct classification versus LDA 1. More Intermediate bifaces were reclassified as both Expert and Novice, and greater numbers of Expert bifaces were reclassified as Intermediate Draft, which more closely resembled Novice than Intermediate handaxes. Even when the omission of

Intermediate Drafts increased overall classification accuracy for LDA 2, the inclusion of Profile Asymmetry decreased the classification accuracy of Intermediates and Novices compared to LDA 1 (excluding Intermediate Drafts).

LDA 3: Adding Flake Scar number

Sample including Intermediate Drafts (n = 130)

- 53.85% jackknifed classification accuracy.

Loadings

Flake Scar number was the weakest discriminator of the three LDA's presented so far across all axes. However, with the addition of Flake Scar number, Axis 1 of LDA 3 explained less variance than Axis 1 for both LDAs 1 and 2 when Intermediate Drafts were included. This result, when combined with the increased jackknifed classification accuracy of LDA 3 versus LDAs 1 and 2, suggests that Flake Scar Number had a positive effect on classification accuracy. This is borne out in greater detail below.

	Axis 1 (73.7%)	**Axis 2 (21.43%)**	**Axis 3 (4.87%)**
Refinement	2.3041	6.1688	0.60567
Edge Straightness	-14.984	16.475	20.455
Planform Asymmetry	-1.1258	7.0787	-4.9142
Th/L	-7.601	-9.9817	-17.026

Flake Scar number	0.061145	0.073801	-0.04186

Table 7. 9. LDA 3: Skill criteria loadings (including Intermediate Drafts).

Confusion Matrix

Compared to LDAs 1 and 2 (including Intermediate Drafts), the inclusion of Flake Scar number in LDA 3 resulted in more Experts and Novices being correctly classified versus LDA 1 and 2 when Intermediate Drafts were included. Additionally, no Experts were classified as Novice, and fewer were misclassified as Intermediate. Conversely, fewer Intermediates were accurately classified than LDA 1, but more than LDA 2.

	Novice	**Intermediate Draft**	**Intermediate**	**Expert**	**Total**
Novice (n=47)	20	12	9	6	47
Intermediate Draft (n=25)	3	17	4	1	25
Intermediate (n=20)	1	2	12	5	20
Expert (n=38)	0	5	12	21	38
Total (n=130)	47	25	20	38	130

Table 7. 10. LDA 3: Confusion Matrix (including Intermediate Drafts).

Scatter plot

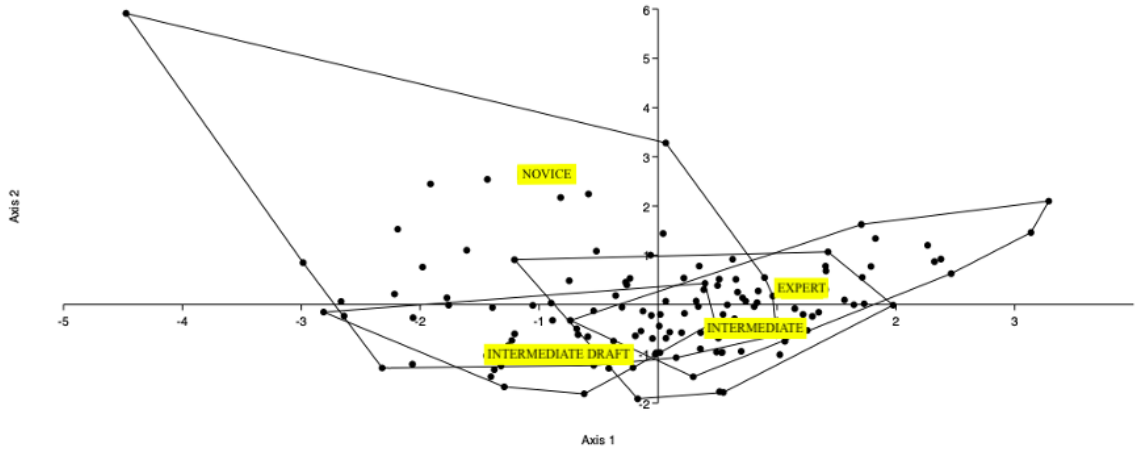


Figure 7. 8. Scatter plot of LDA 3 jackknifed experimental biface scores grouped by reclassified skill-level (including Intermediate Drafts).

Sample excluding Intermediate Drafts (n = 105)

- 60.95% jackknifed classification accuracy.

Loadings

Flake Scar number again contributed the least to both axes.

	Axis 1 (90.28%)	**Axis 2 (9.72%)**
Refinement	1.551	2.4757
Edge Straightness	-19.147	-7.321
Planform Asymmetry	-3.2373	7.4678
Th/L	-5.1494	7.3073
Flake Scar number	0.037294	0.082262

Table 7. 11. LDA 3: Skill criteria loadings (excluding Intermediate Drafts).

Confusion Matrix

When Intermediate Drafts were discounted, the addition of Flake Scar number in LDA 3 increased the classification accuracy of all three skill-levels, especially Expert and Novice, versus LDA 2. Additionally, compared to LDA 1 (excluding Intermediate Drafts), the insertion of Flake Scar number improved the classification accuracy of Experts, matched it for Intermediates, and only slightly lowered it for Novices.

	Novice	**Intermediate**	**Expert**	**Total**
Novice (n=47)	29	12	6	47
Intermediate (n=20)	2	13	5	20
Expert (n=38)	3	13	22	38
Total (n=105)	34	38	33	105

Table 7. 12. LDA 3: Confusion Matrix (excluding Intermediate Drafts).

Scatter plot

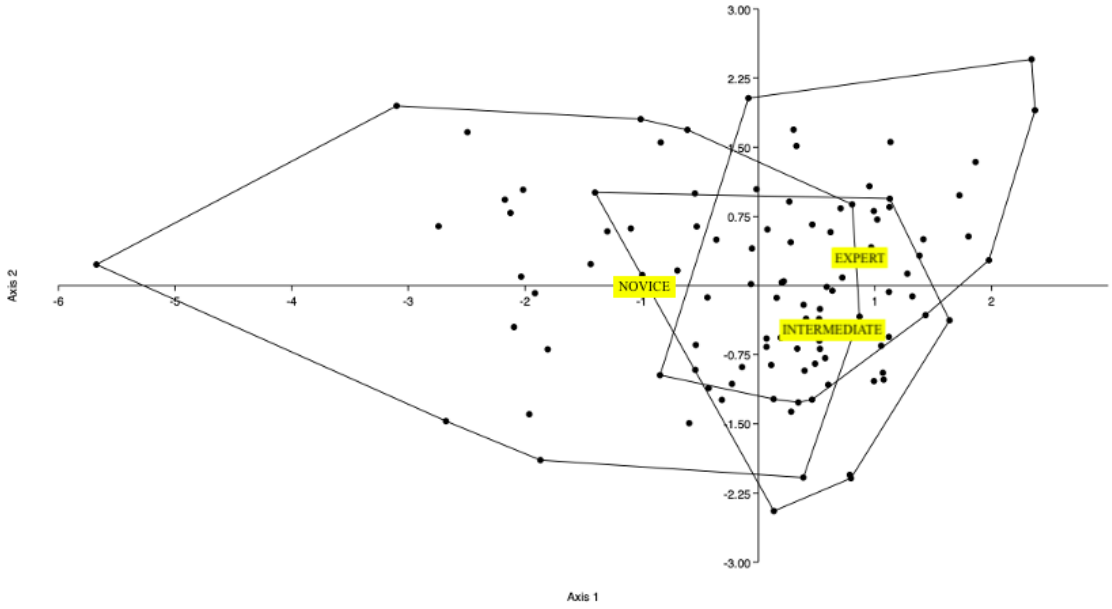


Figure 7. 9. Scatter plot of LDA 3 jackknifed experimental biface scores grouped by reclassified skill-level (excluding Intermediate Drafts).

Again, both samples (including and excluding Intermediate Drafts) showed that LDA 3 was good for distinguishing between Expert/Intermediate and Novice, but not as suitable for distinguishing between Expert and Intermediate.

LDA 4: Adding Flake Scar Number & Profile Asymmetry

Sample including Intermediate Drafts (n = 130)

- 52.31% jackknifed classification accuracy.

Loadings

Overall, Flake Scar number contributed less variance than Profile Asymmetry across all three axes (when Intermediate Drafts were included). Axis 1 explained less variance than

LDAs 1 and 2, but more than LDA 3. This suggested that the addition of both Flake Scar number and Profile Asymmetry had a positive effect on classification accuracy compared to LDAs 1 and 2, but not as much as using just Flake Scar number alone in LDA 3. This is supported by a comparison of jackknifed classification accuracies and is substantiated in greater detail below.

	Axis 1 (74.54%)	**Axis 2 (20.45%)**	**Axis 3 (5.01%)**
Refinement	2.99	6.3046	0.98989
Edge Straightness	-15.353	16.646	19.108
Planform Asymmetry	-0.93231	7.092	-4.7029
Th/L	-8.6831	-10.207	-16.729
Flake Scar Number	0.051988	0.069263	-0.055
Profile Asymmetry	-2.0769	-0.63379	-2.2282

Table 7. 13. LDA 4: Skill criteria loadings (including Intermediate Drafts).

Confusion Matrix

When Intermediate Drafts were included the combination of both Flake Scar number and Profile Asymmetry in LDA 4 slightly reduced classification accuracy versus Flake Scar number alone (LDA 3) for all skill-levels apart from Intermediate Drafts. It also resulted in greater classification accuracy among Experts, Intermediate Drafts, and Novices than using Profile Asymmetry alone (LDA 2), whereas Intermediate classification accuracy remained equal to that in LDA 2. Additionally, more Experts, Intermediates Drafts and

Novices were accurately classified versus LDA 1, but 3 fewer Intermediates were classified correctly.

	Novice	Intermediate Draft	Intermediate	Expert	Total
Novice (n=47)	19	11	10	7	47
Intermediate Draft (n=25)	4	18	2	1	25
Intermediate (n=20)	1	2	11	6	20
Expert (n=38)	0	6	12	20	38
Total (n=130)	24	37	35	34	130

Table 7. 14. Table 7.14. LDA 4: Confusion Matrix (including Intermediate Drafts).

Scatter plot

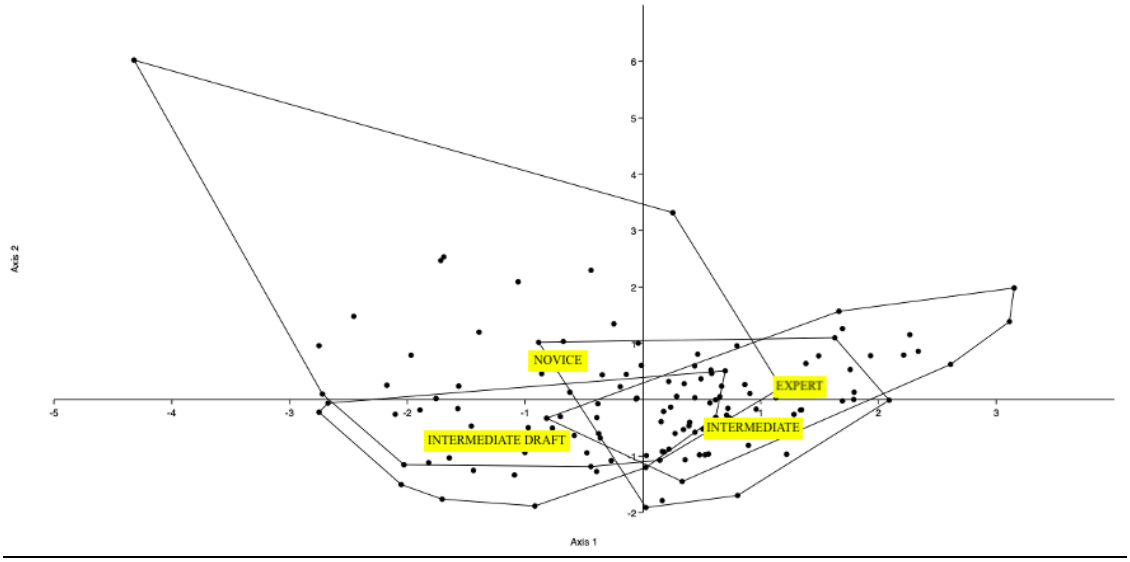


Figure 7. 10. Scatter plot of LDA 4 jackknifed experimental biface scores grouped by reclassified skill-level (including Intermediate Drafts).

Sample excluding Intermediate Drafts (n = 105)

- 58.1% jackknifed classification accuracy.

Loadings

	Axis 1 (90.73%)	**Axis 2 (9.27%)**
Refinement	2.2257	2.1678
Edge Straightness	-19.553	-7.4456
Plan Asymmetry	-2.9193	7.1821
Th/L	-6.2686	7.656
Flake Scar Number	0.029569	0.085828
Profile Asymmetry	-2.2211	1.0375

Table 7. 15. LDA 4: Skill criteria loadings (excluding Intermediate Drafts).

Confusion Matrix

When excluding Intermediate Drafts for LDA 4, correct classification of Experts increased versus LDAs 1, 2, and 3, but for other skill-levels classificatory accuracy decreased versus LDAs 1 and 3 and was equal to that of LDA 2.

	Novice	Intermediate	Expert	Total
Novice (n=47)	26	12	9	47
Intermediate (n=20)	2	12	6	20
Expert (n=38)	3	12	23	38
Total (n=105)	31	36	38	105

Table 7. 16. LDA 4: Confusion Matrix (excluding Intermediate Drafts).

Scatter plot

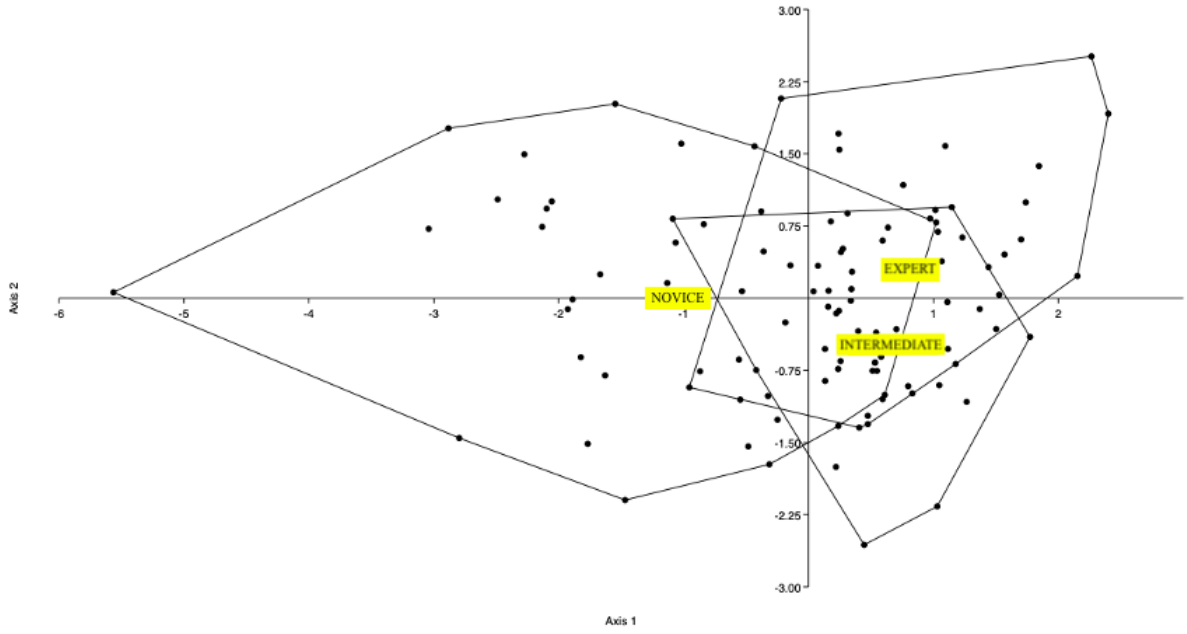


Figure 7. 11. Scatter plot of LDA 4 jackknifed experimental biface scores grouped by reclassified skill-level (excluding Intermediate Drafts).

The addition of both Profile Asymmetry and Flake Scar number achieved the same overall jackknifed classification accuracy as LDA 1 (excluding Intermediate Drafts) and achieved higher classification accuracy than LDA 1 when including Intermediate Drafts (see Table 7.17). However, doing so over-classified Experts versus other skill-levels due mainly to the ambiguity derived from Profile Asymmetry scores disclosed in the univariate and bivariate analyses above. In fact, using Flake Scar number alone attained the highest classification percentages whether Intermediate Drafts were included or excluded. A summary of the results is presented in the below table (Table 7.17).

Quant skill combination	***n = 105 (excluding Intermediate Drafts)***	***n = 130 (including Intermediate Drafts)***
LDA 1	58.1%	50%
LDA 2	53.3%	47.7%
LDA 3	60.95%	53.85%
LDA 4	58.1%	52.3%

Table 7. 17. Summary of LDA accuracy (including and excluding Intermediate Drafts).

In sum, the four LDAs presented above (including and excluding Intermediate Drafts) showed that Discriminant analysis was useful for distinguishing between higher and lower skill-levels when using the quantitative criteria outlined. However, as suggested at the start of this section, there were overlaps between the skill-levels that remained - predominantly between Intermediate and Expert. LDA also confirmed the inefficacy of Profile Asymmetry for discriminating between skill-levels.

However, because my experimental data was going to be used to classify predominantly ovate archaeological bifaces from GZnivG, it was necessary to additionally evaluate the influence, previously shown in my bivariate analysis, that biface shape had on classification accuracy. Thus, in the next section (7.2.1.4), I performed the same LDAs as above but separately for ovate and pointed experimental bifaces. I did this to examine whether my quantiative criteria were better at correctly classifying one biface shape over another, and to make sure that the most accurate LDA data was used on the GZnivG assemblage.

7.2.1.4 Influence of shape

I used Roe's (1968) metric distinctions for point, ovate, and cleaver in my analyses. However, before presenting the shape-sorted discriminant analysis I provide univariate statistics for each skill-level separated by biface shape.

7.2.1.4.1 Univariate results

For the main skill-related criteria, the univariate statistics demonstrated that shape did affect the range of scores generated.

Pointed bifaces exhibited larger ranges for both refinement ratios (see Catalogue 4, Sheet 4.23, Tables 6 and 7) and subsumed more variation in scores than ovates for both Intermediates and Experts. The most obvious reason for this dissimilarity was that points

were made with and without heavier butts, whereas ovates generally never exhibited thicker bases. However, Novice refinement scores were comparable regardless of shape. This underlines the fact that Novices were less in control of the shape and dimensions they intended to produce.

The results also showed that for all skill levels pointed bifaces were less symmetrical than ovates and demonstrated a greater range of symmetries (see Catalogue 4, Sheet 4.23, Table 8).

On the other hand, Edge Straightness was not overly affected by shape. In fact, Edge Straightness data reinforced the view that this was an effective criterion for separating Novices from Intermediate and Expert handaxes, regardless of biface shape (see Catalogue 4, Sheet 4.23, Table 10).

Overall, therefore, even when adjusted for biface shape, univariate scores for each skill-related variable still revealed that the mean scores of Expert and Intermediate ovate bifaces did not differ by more than each groups' standard deviation/variance. Nevertheless, pointed bifaces generally produced a more inconstant range of scores that overlapped more frequently among all skill-level categories. This suggests that using experimental ovate data would be far superior to using mixed (ovate and pointed) experimental data in discriminating between skill-levels among an ovate archaeological assemblage like GZnivG.

Indeed, when bivariate correlations were run again on the main skill-related criteria for only the ovate bifaces (see Catalogue 4, Sheets 4.10, 4.14, 4.18, and 4.22), they showed a general improvement between both Refinement ratios and Edge straightness

correlations and Planform Asymmetry and Edge straightness correlations. However, correlations between Refinement ratios and Planform Asymmetry generally decreased in strength and significance once pointed bifaces were removed from the analysis. Therefore, to further ascertain the effect shape had on the discriminative accuracy of my skill-related criteria, it was necessary to perform the same LDA combinations (1 - 4) as above but split by ovate and pointed bifaces.

7.2.1.4.2 LDA via shape

Each LDA analysis, whether for the entire assemblage, just the ovates, or only the points, tested the classification accuracy when Intermediate Drafts were included and excluded. The percentage accuracy (jackknifed) for each LDA is outlined in the below table.

LDA	***Total sample n = 105 (Excl. Intermed Drafts***	***Total sample n = 130 (Incl. Intermed Drafts***	***Ovate n = 56 (Excl. Intermed Drafts)***	***Ovate n = 71 (Incl. Intermed Drafts***	***Pointed n = 46 (Excl. Intermed Drafts***	***Pointed n = 54 (Incl. Intermed Drafts***
LDA1	58.1%	50%	58.2 %	57.75%	54.35%	40.74%
LDA2	53.3%	47.7%	57.14%	50.7%	58.7%	46.3%
LDA3	60.95%	53.85%	58.93%	54.93%	63.04%	50%
LDA4	58.1%	52.3%	62.5%	54.93%	63.04%	48.15%

Table 7. 18. Summary of LDA accuracy.

Table 7.18 reveals that skill criteria were more effective at discriminating between ovate shape handaxes than pointed ones. This is advantageous given the archaeological assemblage of GZnivG is overwhelmingly ovate.

Profile Asymmetry was again shown to be ambiguous and unrelated to skill-level in this sample and therefore was not used for distinguishing skill-levels in the archaeological assemblage. Furthermore, although Flake Scar number was a better indicator of skill-level, I was unable to record and apply these data because of Covid-19 travel restrictions.

For all LDA combinations, the inclusion of Intermediate Drafts reduced the classification accuracy. However, one must consider in the archaeological record that some bifaces were not made as people wanted and may therefore have been discarded. Indeed, as Questionnaire participant 7 (see Catalogue 2, Q13) perceptively noted, archaeological assemblages must also have included unfinished bifaces. Thus, keeping the Intermediate Drafts as part of the discriminatory process was a useful check/balance when using the experimental data to analyse archaeological data.

Of all the LDA combinations when Intermediate Drafts were included, LDA 1 (which did not include Profile Asymmetry or Flake Scar number) was the most accurate (57.75%) at classifying experimental ovates (n=71). The results are presented below and will be used to directly analyse the ovate biface assemblage from GZnivG (previously introduced in section 7.4).

Loadings

Edge Straightness and Thickness/Length ratio explained most of the variance in both Axes 1 and 2.

	Axis 1 (73.72%)	Axis 2 (22.83%)	Axis 3 (3.45%)
Refinement	-4.4787	11.823	-7.8905
Edge Straightness	13.529	26.062	26.224
Planform Asymmetry	6.0463	-5.1279	4.9291
Th/L	16.796	-22.277	-5.7926

Table 7. 19. LDA 1 ovate only: Skill criteria loadings (including Intermediate Drafts).

Confusion Matrix

- 58% jackknifed classification accuracy.

The breakdown by skill-level is presented in Table 7.20. The key findings were:

- Expert bifaces were never misclassified as Novice items.
- 70% of Intermediate handaxes were accurately classified.
- Only three Novice bifaces were misclassified as Expert.
- Intermediate Drafts were 73% accurately classified.

Almost inevitably some overlapping of categories occurred, but this mainly involved adjacent skill-levels. This could imply the existence of thresholds between skill-levels along a spectrum that relates to Gray and Lindestedt's concept of PDLs: from Novice to Intermediate and vice versa, and from Intermediate to Expert and vice versa. Indeed, intersections may be related to both Novices and Intermediates improving with practice

over time and/or occasionally receiving tuition. Moreover, they could be applied to Experts when their performance occasionally dipped. This pattern arguably shows that the quantitative criteria were more reliable than the 57.75% classification accuracy actually suggests. But there are clearly exceptions to the rule as in the 3 Novice handaxes misclassified as Expert examples. Here it is interesting to note that they were all made by Experimental knapper 4 in their first year of knapping and potentially indicate innate skills in some individuals.

Additional detail explaining the reasons for quantitative misclassifications like those outlined for Experimental knapper 4 is provided by qualitative analysis of the experimental assemblage in section 7.2.2. Nevertheless, for the most part, the misclassifications were not grossly wrong. Instead, they highlight the blurred boundary between skill-levels adjacent to one another and in terms of experience.

	Novice	**Intermediate Drafts**	**Intermediate**	**Expert**	**Total**
Novice	12	6	6	3	27
Intermediate Drafts	1	11	1	2	15
Intermediate	1	1	7	1	10
Expert	0	1	7	11	19
Total	14	19	21	17	71

Table 7. 20. LDA 1 ovate only: Confusion matrix (including Intermediate Drafts).

Scatter plot

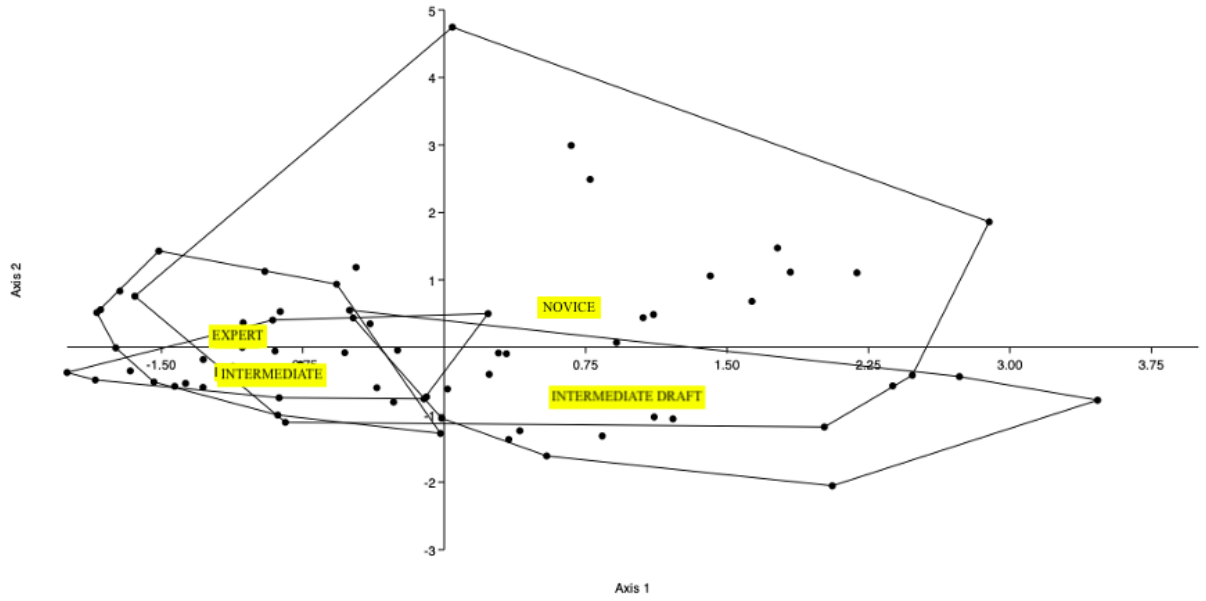


Figure 7. 12. Scatter plot of LDA 1 (ovates only) jackknifed experimental biface scores grouped by reclassified skill-level (including Intermediate Drafts).

7.2.1.5 Influence of practice and Instruction

Separately, as outlined in my Methodology, I studied the influence of practice and instruction on one Novice, as well as the effect of instruction by an Expert on one Intermediate.

The Novice (me/Experimental knapper 7) produced bifaces over 6 months, starting with a basic practical knowledge of handaxe manufacture. This collection of bifaces was added to after each knapping session. The collection was stopped at six months due to the Covid-19 pandemic and my subsequent inability to source flint, as well as a lack of space to knap regularly. On the other hand, the Intermediate (Experimental knapper 4) was a DPhil student who, as part of their research, regularly made bifaces. They stored and kept record

of when they made them, which allowed them to be placed in chronological order for purposes of tracking development.

The Novice (me) received tuition three times during this investigation, producing 9 bifaces. Bifaces were also made as part of a knapping workshop where I observed other knappers of varying skill-levels. One of the 5 bifaces produced during the workshop session was completed with advice, but not direct instruction, from an Expert. As mentioned in my Methodology, advice differs from instruction in the specificity of the information imparted. Here, advice is qualified as more generalised, whereas instruction is highly specific and focused. Overall, I contributed 33 bifaces, 9 under instruction, 1 with advice, and 23 without direct instruction.

The Intermediate's sample from all years of practice consisted of bifaces made without tuition, apart from one ('withKL'). This one was made under step-by-step instruction from an Expert. The groups corresponded to bifaces made in their first year (1yrACCEP), drafted pieces made in their second year (2yrDRAFT'), and those made in their third year of learning after Expert instruction ('postKL').

I will first report observations of the univariate statistics by individual skill-related criteria gathered from the Novice bifaces. Incidentally, the rows for both Novice and Intermediate equate to periods of knapping time that can be grouped together by individual knapping session and whether the bifaces comprising that group were made with or without tuition. The groups are presented in rows in chronological order of manufacture in Catalogue 4, Sheet 4.24, Tables 1 – 6 (for Novice Experimental knapper 7) and in Catalogue 4, Sheet

4.25, Tables 1 – 6 (for Intermediate Experimental knapper 4). The purpose of this is to demonstrate whether there has been any development over time or by instruction.

Following this, I present an LDA which looked at whether multiple criteria can accurately predict whether a biface was made under instruction. My aim was to evaluate whether instruction aided skill development, in quantitative terms, over time when the Novice was subsequently manufacturing bifaces without tuition. I follow the same format for the Intermediate, but the LDA was here mainly focused on development of skill apart from instruction.

7.2.1.5.1 Novice (Experimental knapper 7) – Univariate results

The univariate statistics, presented in Catalogue 4, Sheet 4.24, revealed that for the main four quantitative skill-related criteria (Ratio of Refinement, Thickness/Length ratio, Planform Asymmetry, and Edge straightness) tuition had a significantly greater instant influence on performance than practice did. Interestingly, however, the data showed that when the Novice made bifaces uninstructed between tuition sessions, they reverted to performing quantitatively similarly to the way they performed before tuition. In the following section, I examine in more detail these apparent leaps and dips in Novice performance via LDA.

7.2.1.5.2 Novice (Experimental knapper 7) – LDA results

Loadings

In Axis 1, which explained most of the variance, the strongest discriminator was Thickness/Length ratio, and the weakest discriminator was Planform Asymmetry.

	Axis 1 (71.52%)	**Axis 2 (23.38%)**	**Axis 3 (4.44%)**	**Axis 4 (0.66%)**
Refinement	4.6801	5.8783	-7.1975	-0.13759
Edge Straightness	-5.7312	-13.76	1.8622	19.112
Planform Asymmetry	2.3389	10.646	7.1414	4.0894
Th/L	10.241	-8.3821	12.862	-6.9588

Table 7. 21. LDA 1 skill criteria loadings for Experimental Knapper 7's assemblage.

Confusion Matrix

When assessing accuracy based on separate sessions of knapping practice, the jackknifed accuracy was only 18%. However, if separate sessions were combined to reflect whether they involved tuition or not, the accuracy improved to 85%. Of the 9 Novice handaxes made with tuition, 7 (78%) were correctly classified, and of the 24 handaxes made without tuition, 88% were correctly classified.

		LH_KL	**LH_JL1**	**LH_JL2**	**LH_PRE**	**LH_PSTKL**	**BUTSER**	**>BUTSER**	**Total**
Tuition	**LH_KL**	0	2	0	0	0	0	0	2
	LH_JL1	1	0	1	0	0	0	0	2
	LH_JL2	1	1	1	1	0	0	1	5
No Tuition	**LH_PRE**	0	0	0	0	3	0	2	5

LH_PSTKL	0	0	2	1	3	1	0	7
BUTSER	0	0	0	1	1	0	3	5
>BUTSER	0	0	1	1	1	2	2	7
Total	2	3	5	4	8	3	8	33

Table 7. 22. Confusion matrix of Experiment knapper 7's bifaces.

Scatter plot

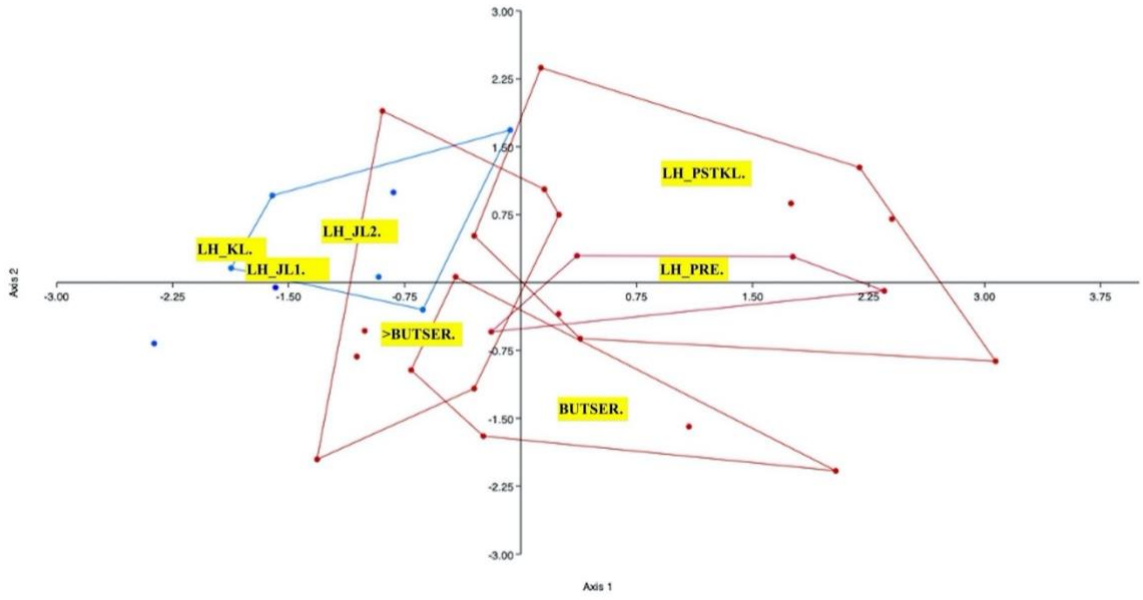


Figure 7. 13. Scatter plot of Experimental knapper 7's jackknifed scores grouped by reclassified session ID.

The scatter plot (Figure 7.13) demonstrates that tuition positively affected the quantitative scores the Novice attained. The blue convex hull (and outlying data points) delineates the bifaces made during Expert tuition. The outlying data points reflect two Expert tuition sessions where only two handaxes were manufactured (there needs to be at least three

bifaces to make a convex hull). The four red convex hulls correspond to bifaces made without tuition.

The only red convex hull that intersects with the blue convex hull representing Expert tuition was the '>Butser', or post-Butser, handaxes. These were the final bifaces made without tuition. However, they were made on a knapping retreat alongside other knappers at Butser. Also, they were made after previously completing two sets of Expert tuition. The closeness of post-Butser, and some Butser handaxes to the bifaces made under Expert instruction, hint that a small skills leap was made by the Novice because of repeated practice and observation of other knappers while at Butser and trying out new observed methods. In addition, some bifaces in the post Expert tuition group (LH_PSTKL), that were made prior to Butser handaxes, also demonstrated that improvments can be made in uninstructed performance after instruction has occurred. But these leaps were inconsistently achieved. In the 'LH_PSTKL' group, there are two handaxes made after Expert instruction that were misclassified as bifaces made under Expert instruction (see Table 7.22). However, most of that group resembled more the bifaces made before Expert instruction. This highlights the leaps and dips associated with learning (Gray and Lindstedt, 2017) and supports my personal account of learning outlined in Chapter 4. Overall, this LDA indicated that the criteria used can accurately predict when tuition had taken place.

7.2.1.5.3 Intermediate – Univariate results

The univariate results, presented in Catalogue 4, Sheet 4.25, showed that, with practice, the Intermediate improved their mean Refinement, Thickness/Length ratio, Edge

Straightness, and Profile Asymmetry scores. However, their Planform Asymmetry and the number of flakes they removed did not change. Moreover, the data suggests that the Intermediate employed more control over thinning, with the standard deviations for Refinement, Thickness/Length ratio, and Profile Asymmetry reducing over time. Yet, Edge Straightness, Flake Scar number, and Planform Asymmetry variance demonstrated no appreciable improvements over time.

However, with Expert instruction, the Intermediate achieved the highest Refinement score and the most flake scar removals of any biface they contributed. Additionally, they achieved the second highest Edge straightness and Thickness/Length ratio of any biface they contributed; only exceeded by the handaxe they made post Expert instruction. However, they achieved worse planform and profile symmetries. This suggests symmetry perhaps held a secondary importance to the Expert instructing them and that focusing on thinning and making straight edges was more important. Indeed, as multiple Questionnaire participants (see responses to Q10, Catalogue 2) have noted, symmetry is secondary to making a functional biface.

7.2.1.5.4 Intermediate – LDA results

The handaxe made by the Intermediate under Expert instruction was not included in my LDA analysis. This is because it did not represent a large enough sample in this category. As such, I could not be certain that the data generated under instruction was reliable, given that I did not have any additional data regarding how variably the Intermediate would perform under Expert instruction.

Two LDAs are presented below: one including the drafted bifaces made in the Intermediate's second year of practice, and one excluding them. To reiterate, the '1yrACCEP' bifaces correspond to the bifaces made by the Intermediate in their first year when they possessed minimal experience and, for the reasons outlined in my Methodology, they were technically a Novice. The '2yrDRAFT' bifaces relate to bifaces ultimately designated as drafts made in their second year of practice while the participant was now, as outlined in my Methodology, considered an Intermediate. Lastly, the 'PostKL' bifaces equate to the bifaces made by the Intermediate in their third year of learning, after they had received one session of Expert tuition.

Including 2yrDRAFTs

- 69.57% jackknifed classification accuracy.

Loadings

The strongest discriminator in Axis 1, which explained most of the variance, was Edge Straightness, and the weakest discriminator was Refinement ratio.

	Axis 1 (80.7%)	**Axis 2 (19.3%)**
Refinement	1.3219	-15.27
Edge Straightness	15.463	-10.672
Planform Asymmetry	9.7076	17.205
Th/L	12.216	9.2193

Table 7. 23. LDA 1 skill criteria loadings for Experimental Knapper 4's assemblage (including Intermediate Drafts).

Confusion matrix

The key outcomes were:

- Non-draft bifaces made in Experimental knapper 4's first year of practice (1yrACCEP), when they possessed minimal experience, were 56% correctly classified.
- Draft bifaces made in Experimental knapper 4's second year of learning (2yrDRAFT), when they, in line with my Methodological classification boundaries, were now categorised as Intermediate, are 72% correctly classified.
- Non-draft bifaces made after Expert tuition in Experimental knapper 4's third year of training (PostKL), when they were still considered, Intermediate, are 75% correctly classified.

	1yrACCEP	**2yrDRAFT**	**PostKL**	**Pre-LDA Total**
1yrACCEP	5	1	3	9
2yrDRAFT	6	18	1	25
PostKL	2	1	9	12
Post-LDA Total	13	20	13	46

Table 7. 24. Confusion matrix of Experiment knapper 4's bifaces (including Intermediate Drafts).

Excluding 2yrDRAFTs

- 66.67% jackknifed classification accuracy.

Loadings

The strongest discriminator was Edge straightness, and the weakest was Thickness/Length ratio.

	Axis 1 (100%)
Refinement	6.4488
Edge Straightness	48.605
Planform Asymmetry	-21.911
Th/L	1.3145

Table 7. 25. LDA 1 skill criteria loadings by Axis for Experimental Knapper 4's assemblage (excluding Intermediate Drafts).

Confusion Matrix

The key findings were:

- The classification accuracy for both non-draft bifaces made in Experimental knapper 4's first and third years of practice ('1yrACCEP' and 'PostKL') did not change with the exclusion of draft bifaces made in their second year of practice ('2yrDRAFT').
- The year 1 and 3 non-draft bifaces (1 each) that were incorrectly classified as year 2 draft bifaces in the above LDA (see Table 7.24), were still misclassified when year 2 draft bifaces were excluded from the LDA (see Table 7.26).

	1yrACCEP	PostKL	Pre-LDA Total
1yrACCEP	5	4	9
PostKL	3	9	12
Post-LDA Total	8	13	21

Table 7. 26. Confusion matrix of Experiment knapper 4's bifaces (excluding Intermediate Drafts).

In sum, the Intermediate's performance improved after Expert instruction, suggesting tuition had an effect, but another factor was greater accumulated practice. Expert instruction, therefore, arguably paved the way to league-stepping methods (Gray and Lindstedt, 2017) that were developed with further self-guided practice. Repeated practice also improved the quantitative scores the Intermediate achieved. As a result, handaxes made in the Intermediate's third year of practice can generally be distinguished from handaxes made during their first year. However, 33% of bifaces made in their first year were misclassified as handaxes made in their third year (after Expert instruction). As per Gray and Lindstedt's (2017) PDLs concept which explains that skill development is liable to fluctuations and unsteadiness in performance, these results may be indicative of leaps and dips in the Intermediate's output.

To further examine the quantitative results of my experimental investigation presented above, I present my qualitative analysis of the bifaces that came from it below. Qualitative assessment was not used here to ascribe a skill-level category to experimental bifaces. Skill-levels were already assigned via methodological separation based on known experience. Instead, qualitative assessment indicated where quantitative analysis could

be improved, especially considering the considerable degree of overlap between Intermediate and Expert quantitative scores.

7.2.2 Qualitative Analysis

7.2.2.1 Results

My sample for qualitative assessment comprised the 30 bifaces that were misclassified as part of the ovate sample of 71 bifaces used for multivariate analysis above (see Table 7.27). The qualitative criteria used to provide context to these misclassifications are outlined in section 6.4.2 of my Methodology. Below, I describe why observations of this kind are worth exploring, as well as assessments of each biface and why they were misclassified.

Point	Given group	Jackknifed	Correct/Incorrect
KL_1.1	Expert	Expert	✓
KL_1.2	Expert	Expert	✓
KL_1.3	Expert	Intermediate	×
JL_1.5	Expert	Expert	✓
JL_1.7	Expert	Intermediate	×
DK2	Expert	Expert	✓
KL_LIV.1	Expert	Expert	✓
KL_LIV.2	Expert	Expert	✓

KL_LIV.3	Expert	Intermediate	×
KL_LIV.7	Expert	Intermediate	×
KL_LIV.8	Expert	Intermediate	×
KL_LIV.12	Expert	Expert	✓
KL_LIV.13	Expert	Expert	✓
KL_LIV.14	Expert	Expert	✓
KL_LIV.15	Expert	Intermediate	×
KL_LIV.17	Expert	Expert	✓
KL_LIV.19	Expert	Intermediate	×
KL_LIV.20	Expert	Intermediate Draft	×
KL_LIV.22	Expert	Expert	✓
CS_LIV1.1	Novice	Expert	×
CS_LIV1.4	Novice	Expert	×
CS_LIV1.5	Novice	Intermediate	×
CS_LIV1.6	Novice	Expert	×
CS_LIV1.8	Novice	Intermediate Draft	×
CS_LIV1.9	Novice	Intermediate	×
CS_LIV2.2	Intermediate	Intermediate Draft	×
CS_LIV2.3	Intermediate	Expert	×
CS_LIV2.5	Intermediate	Intermediate	✓
CS_LIV2.8	Intermediate	Intermediate	✓
CS_LIV2.9	Intermediate	Intermediate	✓
CS_LIV2.10	Intermediate	Intermediate	✓
CS_LIV2.13	Intermediate	Intermediate	✓
WM_PREJL.2	Intermediate	Intermediate	✓
WM_JL.1	Intermediate	Novice	×
WM_JL.5	Intermediate	Intermediate	✓
ADAMTURN.2	Novice	Novice	✓
ADAMTURN.3	Novice	Intermediate Draft	×
ADAMTURN.4	Novice	Intermediate Draft	×

ADAMTURN.5	Novice	Novice	✓
LH_KL.2	Novice	Intermediate	×
LH_JL1.1	Novice	Intermediate	×
LH_PRE.1	Novice	Novice	✓
LH_PRE.3	Novice	Novice	✓
LH_PRE.5	Novice	Novice	✓
LH_PSTKL.1	Novice	Novice	✓
LH_PSTKL.2	Novice	Novice	✓
LH_PSTKL.3	Novice	Novice	✓
LH_PSTKL.6	Novice	Novice	✓
LH_PSTKL.7	Novice	Novice	✓
BUTSER.1	Novice	Intermediate Draft	×
BUTSER.2	Novice	Intermediate Draft	×
BUTSER.3	Novice	Intermediate Draft	×
>BUTSER.1	Novice	Intermediate	×
>BUTSER.2	Novice	Intermediate	×
>BUTSER.5	Novice	Novice	✓
>BUTSER.6	Novice	Novice	✓
CS_LIV3.1	Intermediate Draft	Intermediate Draft	✓
CS_LIV.3	Intermediate Draft	Intermediate Draft	✓
CS_LIV3.5	Intermediate Draft	Novice	×
CS_LIV3.6	Intermediate Draft	Intermediate Draft	✓
CS_LIV3.8	Intermediate Draft	Intermediate Draft	✓
CS_LIV3.9	Intermediate Draft	Expert	×
CS_LIV3.11	Intermediate Draft	Intermediate Draft	✓
CS_LIV3.12	Intermediate Draft	Expert	×
CS_LIV3.13	Intermediate Draft	Intermediate Draft	✓
CS_LIV3.14	Intermediate Draft	Intermediate	×
CS_LIV3.15	Intermediate Draft	Intermediate Draft	✓
CS_LIV3.17	Intermediate Draft	Intermediate Draft	✓

CS_LIV3.18	Intermediate Draft	Intermediate Draft	✓
CS_LIV3.20	Intermediate Draft	Intermediate Draft	✓
CS_LIV3.23	Intermediate Draft	Intermediate Draft	✓

Table 7. 27. Comparison of original and jackknifed classifications for Experimental ovate bifaces.

1. CS_LIV1.1 to CS_LIV1.9 (bifaces made by Experimental knapper 4 in their first year of practice when classified as a Novice)

Observation:

- None of these bifaces were correctly classified as Novice (3 Expert, 2 Intermediate, and 1 Intermediate Draft).

Assessment:

- Compared to other Novices in the sample, Experimental knapper 4 was achieving quantitative scores in their first year of practice equivalent to knappers with more than 10 years greater experience. However, the three bifaces classified as Expert were in some ways qualitatively different to those made by real Experts, despite being more qualitatively similar to Expert bifaces than other Novices in the sample:

- CS_LIV1.1:



Figure 7. 14. Biface CS_LIV1.1 in Planform and Profile views.

This apparently Expert handaxe showed no instances of battering and there were very few failed terminations. However, versus other experimental Expert bifaces, significant invasive flaking was absent. Moreover, the edges were straight but not as sharp as other Expert examples. The biface in general, and the tip specifically, could be thinned better and more evenly – something demonstrated by the others. However, the main qualitative difference was the feel of the edge. Experts produced far sharper edges. On the other hand, this biface displayed some 'advanced features' in the skilled maintenance of the edge and plane of intersection, as well as better thinning and imposition of symmetries. In addition, it exhibited edge regularisation after thinning which is something Experts performed but Novices typically did not.

- CS_LIV1.4:



Figure 7. 15. Biface CS_LIV1.4 in Planform and Profile views.

This biface also showed no instances of battering and very few failed terminations. However, versus other experimental Expert bifaces, significant invasive flaking was less pronounced. The biface was plano-convex in profile due to being made on a large flake. Improved refinement ratios resulting from the reduced thickness of it being made on a flake has helped the biface be quantitatively classified as Expert. The knapper found it hard to move the plane of intersection more centrally between the two faces and, as such, the biface remained planoconvex due to the knapper's poor control of the biface profile. Again, the edges did not feel sharp compared to experimental Expert handaxes.

However, the general symmetry, straight edges, and other behaviours, such as post-thinning edge regularisation, were more like Experts than other Novices in the experimental sample. Consequently, although qualitative assessment did not justify its quantitative classification of Expert, its demonstrated performance levels were higher than bifaces made by other Novices. This biface, made in year 1 of Experimental knapper

4’s learning, more qualitatively resembled other bifaces made by the same knapper in the years following when they were methodologically classified as Intermediate. Such outcomes may have been mitigated and distinctions made clearer if I had been able to incorporate more participants in the study.

- CS_LIV1.6:



Figure 7. 16. Biface CS_LIV1.6 in Planform and Profile views.

This biface showed no instances of battering. However, there were multiple failed terminations. The knapper found it hard to thin the biface, especially at the tip. Again, the edges did not feel sharp compared to experimental Expert handaxes. However, the general symmetry, straight edges, and behaviours such as post-thinning edge regularisation, were more like experimental Experts than other Novices in the sample. Yet, the numerous mistakes visible in the final biface gave the impression that the knapper was not in full control of their actions. Qualitative assessment, therefore, contradicted the quantitative classification given.

Overall, these misclassifications reflect some of the limitations of the quantitative methods to discriminate between skill-levels. However, Experimental knapper 4 consistently qualitatively outperformed other Novice knappers in this investigation and was clearly the exception to the rule. After a year of experience, they had evidently reached Intermediate levels in certain criteria. Qualitative assessment thereby supports the notion that the predominant shortcoming of my quantitative methodology was distinguishing between Intermediate and Expert and that identifying true Novices was relatively straightforward.

2. KL_LIV.20 and CS_LIV3.12

Observation:

- KL_LIV.20 was the only Expert biface to be classified as Intermediate Draft, whilst CS_LIV3.12 was one of two Intermediate Draft bifaces that were misclassified as Expert.

Assessment:

- KL_LIV.20:



Figure 7. 17. Biface KL_LIV.20 in Planform and Profile views.

The smaller size of this biface negatively affected its quantitative classification. Its Th/L ratio score was above the mean score for this criterion for both Novice and Intermediate Draft experimental skill categories. Its Refinement ratio was also between scores typical of Intermediate Draft and Novice and nowhere near that of Intermediates or Experts. However, its Edge Straightness was more like Experts and Intermediates.

Qualitative assessment broadly justified quantitative classification as Intermediate Draft but was ambiguous in other respects. The biface displayed multiple instances of poor thinning where step and hinge fractures were visible, as well as areas of high topography on either face. Additionally, the edge was quite zig-zagged and did not appear to have been regularised much after thinning. This did not look like a particularly skilled handaxe, and its poor execution could have been due to the raw material, although it was not of obvious poor quality. Alternatively, it could have signified a short-lived dip in performance on the part of the knapper. Expediency of the knapper during manufacture was also considered as a potential explanation for the visible mistakes, but this was ultimately disregarded due to the knapper not being placed under any time pressure.

- CS_LIV3.12:



Figure 7. 18. Biface CS_LIV3.12 in Planform and Profile views.

This biface was made on a flake which assisted in reducing the TH/L ratio score more in line with experimental Expert and Intermediate means. Its Edge Straightness was nearer to Expert and Intermediate means too. Scores commensurate with Expert and Intermediate performance in these two criteria, which were the main discriminating loadings of the LDA used to reclassify experimental handaxes (see Table 7.19), influenced this biface's positive misclassification.

Qualitatively, this biface exhibited clear evidence of inexpert performance. Firstly, the knapper repeatedly utilised a hard hammer to remove thinning flakes when the use of soft hammers would have been more helpful for extensive thinning. Indeed, multiple step/hinge fractures indicate thinning was unsuccessful. Furthermore, one edge was much straighter than the other which suggested this knapper also possessed inexpert edge management.

Qualitative comparison implies that the quantitative criteria used were on one hand potentially blind to dips in performance of Experts, and inflexible and dependent on how most of the bifaces in a sample were made. In this experimental investigation, most handaxes were made by reduction of a larger nodule, with only a few being made on flakes. Therefore, the lower refinement ratio scores that Experts achieved more regularly and were key to discrimination using LDA, can also be achieved by lower skill-levels when making a biface on a flake.

3. CS_LIV2.13

Observation:

- This biface was made by an Intermediate under Expert instruction but remained classified as Intermediate and not Expert. This was unexpected given that other bifaces made by this knapper (4) earlier in their development were reclassified as Expert. Whereas most bifaces made after instruction remained classified as Intermediate.

Assessment:

CS_LIV2.13 achieved the lowest ratio of refinement (0.196), second lowest Thickness/Length ratio (0.14), and second lowest Edge Straightness scores (0.041) of Experimental knapper 4's contribution to the experimental investigation. The lowest Th/L ratio and Edge Straightness scores for this knapper came after the Expert instruction that was received during the making of biface CS_LIV2.13. This showed that Expert tuition, as shown above, may have helped to provide a leap in Experimental knapper 4's performance.



Figure 7. 19. Biface CS_LIV2.13 in Planform and Profile views.

Compared to Experts in the ovate sample, this biface displayed a lower refinement and Th/L ratio and was equal to the lowest (i.e., best) Expert score in Edge Straightness. In fact, it scored the lowest score on Axis 1 of the whole sample. This was most probably the consequence of Expert instruction. Indeed, this biface sat at the lower end of the Intermediate range on Axis 1 and below that of Expert (0.102 versus 0.109 for Expert). But if this biface was excluded, the lower end of the Intermediate range on Axis 1 would become equal to that of experimental Experts. This indicated that Experimental knapper 4 could leap in performance with Expert help to a level above that reached by Experts and confuse LDA classification accuracy. Importantly, exposing the power that tuition can have on performance is relevant given that tuition was something that likely occurred in the Acheulean. Therefore, its effect cannot be ignored when interpreting skill-level from the bifaces left in the archaeological record.

Biface CS_LIV2.13 also emphasised how difficult it was to separate this Intermediate knapper, using the quantitative criteria employed, from the three Expert knappers in this investigation. Because this biface scored below (i.e., better) than the Expert bifaces on Axis 1, it may have remained classified as Intermediate because it did not fall within the Expert range (0.109 – 0.189). However, this classification, as per my Methodology, was jackknifed and did not use its own data in its classification. Moreover, most bifaces below the Axis 1 mean (0.14) for Experts were made by Experts (53%), whereas Intermediates constituted only 31% below that mean. Nevertheless, the similarity in Axis 1 scores between the two skill-levels meant that either an Intermediate or Expert classification would have been equally valid.

Qualitatively, however, versus the three Novice bifaces above (CS_LIV1.1, CS_LIV1.4, and CS_LIV1.6) that were reclassified as Expert, this correctly classified Intermediate biface showed much greater control of edge maintenance and thinning. The ability to thin, whilst keeping the biface broad was demonstrated well. This was supported by this biface reaching the best Ratio of refinement score of the entire assemblage. The sharpness of the edge to the touch was also much improved versus others made by Experimental knapper 4 - it even resembled quite closely the Expert bifaces. Again, the effects of tuition here were clear and are thus something to be considered when interpreting skill-level in the archaeological record.

4. WM_JL.1

Observation:

- This biface was made by Experimental knapper 5 and was the only Intermediate biface classified as a Novice.

Assessment:



Figure 7. 20. Biface WM_JL.1 in Planform and Profile views.

Edge Straightness, the second strongest loading affecting LDA classification (see Table 7.19), was the principal quantitative criterion that affected the reclassification of biface WM_JL.1. Experimental knapper 5 attained an Edge Straightness score equivalent to the mean of Novices (0.1) which is more than the maximum Expert score (0.09) and equivalent to the worst score of the Intermediate range.

Qualitative analysis confirmed that Experimental knapper 5's edge maintenance was not as good as Experimental knapper 4 (other Intermediate), and inferior to the other two bifaces they themselves had contributed to the assemblage. A result of Experimental knapper 5's poor edge maintenance was one irregularly shaped, crooked edge in profile. This was exacerbated by the knapper's inability to prepare suitable platforms to thin the

convex face. Additionally, there was no post-thinning regularisation visible, and the knapper was unable to remove all cortex from the biface. Despite this, the plane of intersection was established and looked more like bifaces made by Experimental knapper 7 (a Novice).

Qualitative assessment was supported by quantitative results:

- This Intermediate biface's Edge Straightness score (0.1) was above (i.e., worse than) the range of means (0.08 – 0.09) for bifaces made by Novice Experimental knapper 7 under instruction from an Expert. In fact, this score corresponds to the lower (i.e., better) end of the range of means for bifaces made by Novice Experimental knapper 7 when no tuition took place (0.1 – 0.15) (see Catalogue 4, Sheet 4.24, Table 4).
- This biface's Thickness/Length ratio (0.26) was also above the range of means (0.18 – 0.23) produced by Experimental knapper 7 under Expert instruction. Similar to Edge Straightness, this biface is at the lower (i.e., better) end of the range of Thickness/Length ratio means for bifaces made by Novice Experimental knapper 7 when they were not being coached by Experts (0.26 – 0.35) (see Catalogue 4, Sheet 4.24, Table 2).
- In addition, this biface's Ratio of refinement score (0.37) fell between the range of means (0.3 – 0.35) generated by Experimental knapper 7 under Expert tuition and the range of means (0.38 – 0.64) made by Experimental knapper 7 without direction (see Catalogue 4, Sheet 4.24, Table 1).

This biface straddled the boundary between Intermediate and Novice. Indeed, its normalised Axis 1 score (0.20037881) lies within the same boundary threshold of

Intermediate/Novice that I use to categorise boundary bifaces in the GZnivG assemblage (see Table 7.38).

5. LH_KL.2 and LH_PSTKL.1

Observation:

- Both bifaces were made by Experimental knapper 7 (me), but one (LH_KL.2) was made under direct, step-by-step Expert instruction and the other (LH_PSTKL.1) was made after instruction but without the help of the Expert. The quantitative results (see Catalogue 4, Sheet 4.1) indicate Experimental knapper 7 experienced a dramatic improvement in performance with instruction, but one that could not be sustained without tuition.

Assessment:

The planform and profile images above (Figures 7.21 and 7.22) plainly evidence the disparity in Experimental knapper 7's performance when making these two bifaces. Biface LH_KL.2, when compared to biface LH_PSTKL.1, demonstrated far greater control on behalf of the knapper in every quantitative and qualitative measure. Biface LH_KL.2 is larger, more refined, has much straighter edges, and thinning has been achieved very well. As such, it resembles a clearly recognisable ovate Acheulean biface that is symmetrical in both planform and profile views, and good management of the plane of intersection has resulted in a suitably biconvex profile shape. Whereas biface LH_PSTKL.1 is much smaller, less refined, displays considerably more sinuous edges, and much poorer management of the plane of intersection. Additionally, analysis of flake scars showed that the knapper found it difficult without instruction to remove long, thin

flakes like those visible on LH_KL.2. This resulted in comminution of LH_PSTKL.1 as the knapper (me) tried in vain to replicate what they had just achieved with tuition.



Figure 7. 21. Biface LH_KL.2 in Planform and Profile views.



Figure 7. 22. Biface LH_PSTKL.1 in Planform and Profile views.

The inconsistency of Experimental knapper 7's (my) performance is best explained by the profound but temporary effect direct step-by-step instruction had on my novice ability to make a handaxe. Through tuition I borrowed the Expert's know-how regarding what to do in terms of choice of percussor, platform preparation, angle of strike, and force of blow, but perhaps more importantly, what to refrain from doing. As I noted in Chapter 4, beginners are less able to see the bigger picture regarding the consequences of current actions and how they affect future decisions. As I explained, this can lead to comminution of the bifaces as the novice reduces the dimensions of the biface by removing too much of the edge in search of something they want to make but do not know how to realise. This is visible in biface LH_PSTKL.1, where without coaching I returned to performing as I had done before tuition (see Catalogue 4, Sheet 4.1). Conversely, the opposite is visible in LH_KL.2, where, with expert guidance, I (Experimental knapper 7) was able to circumvent cortex issues at the biface tip and maintain biface length and a sharp tip.

Overall, the comparison of these bifaces demonstrates that if found in the archaeological record, and assuming direct instruction cannot be proved, they would be presumed to have been made by very differently abled knappers. This would have a significant effect on any interpretation of who contributed to the assemblage and what was going on at the site.

7.2.2.2 Summary

The bifaces analysed and discussed in this section constituted interesting examples that demonstrated the ability of qualitative assessment to validate or contest quantitative

classifications. Overall, qualitative assessment was very helpful in determining whether quantitative classification of skill-level was reasonably accurate by providing greater contextual detail.

For example, qualitative observation highlighted that, counter to their quantitative classification, bifaces CS_LIV1.1, CS_LIV1.4, and CS_LIV1.6 generally did not exhibit Expert behaviour. Nonetheless, all 3 of these Novice bifaces exhibited behaviours more proportionate to Intermediate-level knapping, thus vindicating the quantitative outcome that these bifaces were of a higher level than other Novice bifaces. Moreover, qualitative assessment of Expert biface KL_LIV.20 vindicated its negative quantitative classification through demonstrating the presence of broadly inexpert knapping performance.

Conversely, these examples demonstrate that correct skill-level classification will be more difficult when applying a mixed (quantitative and qualitative) method to archaeological bifaces where the skill-level of the knapper is unknown. For instance, examples such as Expert biface KL_LIV.20, Intermediate biface CS_LIV2.13, and Novice bifaces LH_KL.2 and LH_PSTKL.1, showed that quantitative declines and/or improvements in performance can be mirrored in qualitative analysis. In other words, Experts can dip in performance and, alternatively, those with less experience have the potential to (especially with tuition) improve dramatically and blur the boundary between the higher and lower skill-levels both quantitatively and qualitatively. Examples made under instruction constitute exceptions in my experimental assemblage, however. Moreover, the step-by-step instruction given to some knappers in my experimental study is a contested method of tuition in the Acheulean and cannot be confidently presumed to have taken place at Gouzeaucourt. I therefore contend that my experimental study has

established that a combination of quantitative and qualitative analysis can reliably differentiate between skill-levels.

Before I present my investigation of skill-level differences at GZnivG, in the next section I present the results of a Blind Test aimed at finding a better way to quantitatively discriminate between Expert and Intermediate bifaces. Given that my original experimental assemblage involved no constraints on biface manufacture, I devised this test to examine whether imposing restrictions on biface manufacture improved quantitative discrimination between the two skill-levels.

7.3 Blind Test

A Blind Test was undertaken on 6 bifaces, 3 of which were made by an Intermediate, the rest

by an Expert. Since both participants had contributed handaxes to the main experimental assemblage, it was necessary to impose morphological constraints on manufacture. This was to reduce the effect of any morphological preference bias that each knapper may have had and that I may have attributed to an individual. Both contributors attempted to replicate an experimentally made biface specimen held by the University of Liverpool. The biface, shown in Figure 6.1, is highly refined and made by a master knapper with multiple decades of experience (John Lord).

Those who participated devised the coding system for the Blind Test themselves. They decided upon names that were attributed to each biface (see Table 7.28). The names of the knappers were kept secret until I had conducted my blind assessment. Based on an LDA analysis I assigned either Expert or Intermediate skill-level classifications to each biface (see Table 7.30). Once I had made my classifications, I checked them with the participants via email.

The sample size of the Blind test was not as big as I would have liked, but due to the complications regarding Covid-19, it was all I was able to organise in the time.

To make my assessments, I normalised the data for all bifaces, including Lord's replica biface specimen, using the results of the experimental LDA presented above in section 7.2.4.2. The standardised values for each criterion were then summed together to give the desired axis score for each specimen (see Table 7.28). Generating normalised axis scores from the experimental assemblage allowed me to identify thresholds on each axis for Novice, Intermediate, and Expert. This was then applied to the Blind test, as well as GZnivG (presented below). The ranges for each skill-level are presented by axis in Table 7.38, in section 7.4.1.2 below.

Biface ID	Axis 1 (73.72%)	Axis 2 (22.83%)	Axis 3 (3.452%)
Rob	0.093	0.021	-0.074
Dave	0.095	0.009	-0.060
Ken	0.131	0.038	-0.098
Arthur	0.156	0.013	-0.096
Fred	0.094	0.021	-0.069

Frank	0.152	0.033	-0.084
John Lord (example)	0.101	0.012	-0.059

Table 7. 28. LDA 1 skill criteria loadings for Blind Test biface assemblage.

The Blind test results all lay within the predicted range of both Expert and Intermediate (see Table 7.38) which showed that both the Expert and Intermediate performed consistently in line with the results they generated in section 7.2. However, there was a clear distinction on Axis 1 between those that were closer to the best performing Expert and Intermediate scores, and those that were closer to the mean scores for both skill-levels. Indeed, specimens Rob, Dave, and Fred were much closer to the Expert replica sample they were copying than specimens Ken, Arthur, and Frank. However, to make this clearer, I performed a PCA on the axes scores for each specimen.

Loadings

	PC 1 (85.86%)	**PC 2 (10.685%)**	**PC 3 (3.455%)**
Axis 1	0.87675	0.41083	0.25005
Axis 2	0.18476	-0.76773	0.61357
Axis 3	-0.44404	0.49175	0.74901

Table 7. 29. PCA loadings of Blind Test bifaces.

In Principal Component (PC) 1, which explains most of the variance, the strongest discriminator was Axis 1 (see Table 7.29). As outlined above in section 7.2.4.1.2, Axis 1 explained most of the variance in the original LDA performed on the experimental ovate assemblage. Th/L ratio and Edge Straightness index were the strongest discriminators within that Axis (see Table 7.19). Therefore, as the original experimental LDA data was

used to normalise the Blind Test data, it follows that Axis 1 was the strongest discriminator in PC1. PC1 is located on the horizontal X axis in the below scatter plot (Figure 7.23), with PC2 being located on the vertical Y axis. PC1 was where the differences in scores between Rob, Dave, and Fred as one group, and Ken, Arthur, and Frank as another, were clear.

Scatter plot

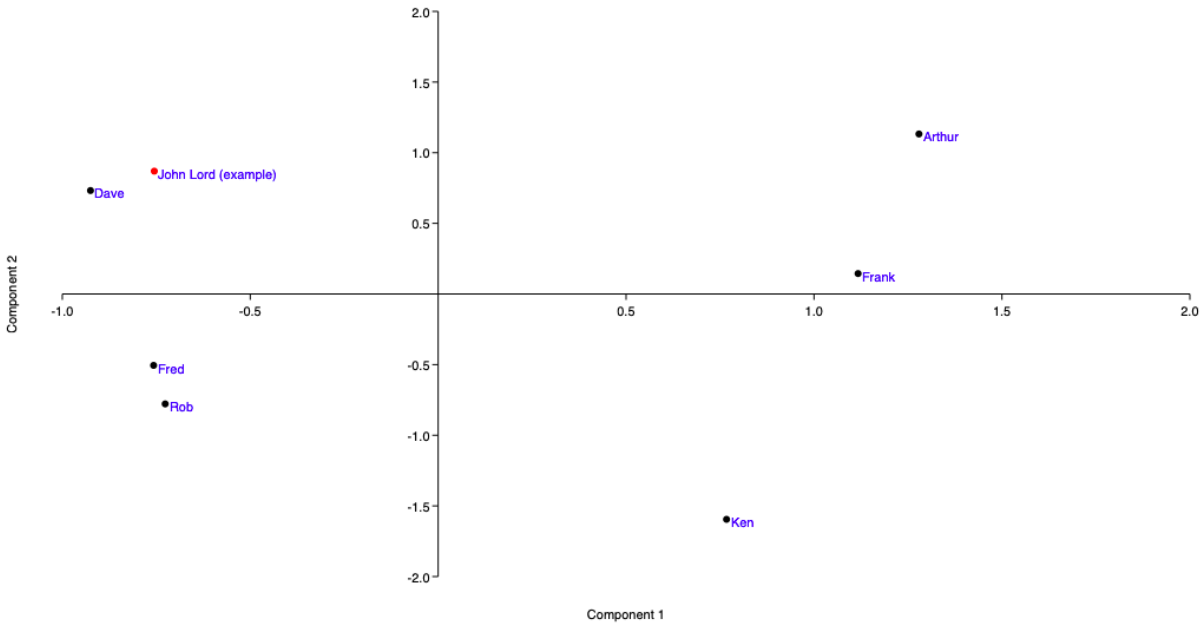


Figure 7. 23. PCA scatter plot of Blind test bifaces and the example biface.

Given the clear distinction between the bifaces, and the reality that there were 3 bifaces produced by both Expert and Intermediate, I submitted the following results for verification:

Biface ID	Skill-level
Rob	Expert
Dave	Expert

Ken	Intermediate
Arthur	Intermediate
Fred	Expert
Frank	Intermediate

Table 7. 30. Skill-level classification given to Blind Test bifaces.

This was confirmed as 100% correct by the one of the participants via email correspondence on 15th November 2021.

The reduced ability of the Intermediate knapper to replicate the example specimen in terms of quantitative criteria is important, especially when compared to the closeness of Intermediate and Expert quantitative scores in the main experimental investigation. Despite the overlap in Intermediate and Expert experimental scores, the range and standard deviation of skill criteria was greater for Intermediates than Experts, suggesting Intermediates performed more variably. This is supported by the Blind Test results, which highlight the increased control, as well as flexibility the Expert knapper had over their engagement with the stone.

Nevertheless, the difference in performance is more marked between the skill-levels in the Blind Test than the principal experimental investigation. Tasking the knappers with replicating a particular morphology had a seemingly significant effect. Interestingly, multiple Questionnaire participants (5, 6, and 8) noted (see Catalogue 2, Q.17) that competence in a knapper can be boiled down to their ability to reproduce a given artefact. Due to the variation in raw material, among other things, replication of a standard requires flexibility, something that Questionnaire participant 1 noted as particularly germane to

the notion of competence. Therefore, imposing restrictions on manufacture may provide a future avenue for separating Expert and Intermediate. Indeed, morphological constraints are relevant to the investigation of skill-level differences because they are visible in the Acheulean record. But admittedly, the sample size of this Blind Test was too small to make more confident claims. To fully test the effect of morphological constraints on performance, this test should ideally be repeated with more participants and a greater sample size.

7.4 Archaeological Investigation

Like the experimental investigation above, a series of quantitative and qualitative analyses is presented. Based on the experimental results, my quantitative examination intends to assign skill-levels to bifaces at Gouzeaucourt. Following this, a qualitative analysis will evaluate the quantitative classifications and address any anomalies.

7.4.1 Quantitative Analysis

Although it cannot be known for certain who made the bifaces at GZnivG, quantitative analysis can be used to discriminate between the better made and more poorly made bifaces as per the criteria and methodology used in the experimental study. Adopting this approach, it should be possible to sub-divide the bifaces according to different skill-sets

and thereby provide a clearer indication of the cross-section of categories of Palaeolithic toolmakers involved.

It is important to reiterate that this study does not attempt to ascertain how many individual knappers contributed to this assemblage. As mentioned previously, this would be impossible without the associated by-products of biface manufacture and extensive refitting. Moreover, GZnivG was not a single-use site, and therefore potentially consists of bifaces from several visits. Consequently, instead of focusing on the Minimum Number of Individual toolmakers present within the assemblage, my assessment of individual bifaces calculates the relative proportions of skill-level categories at GZnivG.

7.4.1.1 Univariate statistics

The details of the quantitative criteria measured from GZnivG bifaces are presented in Catalogue 3, Sheet 3.2. The results will be compared to the experimental investigation data outlined in Catalogue 4, Sheet 4.5.

Length: The experimental sample included a greater range of lengths than in GZnivG. While the minimum scores were similar in both assemblages, in the experimental sample the smaller size of biface can be attributed to lower skill-levels. The same cannot be assumed in the archaeological assemblage where some of the bifaces had undergone re-sharpening (McPherron, 1994; Soriano, 2000).

Thickness: GZnivG bifaces displayed a much smaller range in thickness than the experimental sample. The mean was also much smaller than the experimental sample (30mm). This was likely due to the differences in length, weight, and shape (i.e., there were more pointed bifaces and no re-sharpening in the experimental sample).

Weight: The experimental sample included much heavier bifaces extending beyond the c.300g GZnivG upper limit. However, the lower limit was very similar. This could imply, as Key and Lycett (2017) found, that a minimum weight task-specific threshold existed below which the utility of the bifaces (archaeological and experimental) was considerably reduced. In the experiments, while the lighter bifaces were all made by Novices, all the other skill categories had a minimum weight above that of the mean of GZnivG. This difference between the assemblages is likely explained by evidence of resharpening at GZnivG which was deliberately left out of the experiment. Because of this the relevance of comparisons of weight between the assemblages was considerably reduced.

Ratio of Refinement: The mean was the same in both assemblages, but the range and variance were increased in the experimental sample. Both the lower and upper limits of the experimental collection were beyond GZnivG's (see Table 7.31). However, the experimental Expert range mirrored closely the Gouzeaucourt range, although the Expert mean in the experimental sample was lower.

	Minimum	***Maximum***	***Range***	***Interquartile range***	***Mean***	***Standard deviation***

Total Experimental (n=130)	0.196	0.917	0.721	-	0.387	0.126
Expert Experimental (n=38)	0.242	0.637	0.395	0.097	0.343	0.084
GZnivG (n=129)	0.24	0.68	0.44	0.08	0.39	0.08

Table 7. 31. Comparison of Ratio of refinement univariate statistics between Experimental and GZnivG assemblages.

Thickness/Length ratio: As with the Ratio of Refinement, GZnivG data revealed its range and variance to be more similar to the experimental Expert data, but its mean more closely resembled the mean of the entire Experimental assemblage (see Table 7.32).

	Minimum	***Maximum***	***Range***	***Interquartile range***	***Mean***	***Standard deviation***
Total Experimental (n=130)	0.131	0.509	0.378	-	0.256	0.072
Experimental Expert (n=38)	0.132	0.366	0.234	0.047	0.221	0.05
Gouzeaucourt	0.15	0.44	0.29	0.07	0.27	0.06

Table 7. 32. Comparison of Th/L univariate statistics between Experimental and GZnivG assemblages.

Planform Asymmetry: GZnivG revealed a comparable mean to the experimental assemblage (see Table 7.33).

	Minimum	*Maximum*	*Range*	*Interquartile range*	*Mean*	*Standard deviation*
Total Experimental (n=130)	0.026	0.387	0.361	-	0.110	0.064
Gouzeaucourt (n=129)	0.04	0.32	0.28	0.06	0.10	0.05

Table 7. 33. Comparison of Planform Asymmetry univariate statistics between Experimental and GZnivG assemblages.

Profile Asymmetry: The range, mean, and variance was reduced in the archaeological assemblage compared to the experimental assemblage (see Table 7.34).

	Minimum	*Maximum*	*Range*	*Interquartile range*	*Mean*	*Standard deviation*
Total Experimental (n=130)	0.084	0.65	0.566	-	0.27	0.132
Gouzeaucourt (n=129)	0.07	0.48	0.41	0.10	0.19	0.08

Table 7. 34. Comparison of Profile Asymmetry univariate statistics between Experimental and GZnivG assemblages.

Edge straightness: The mean and minimum scores were similar to the experimental assemblage (see Table 7.35). Although the range was quite different, the greater upper limit of the experimental sample (0.36) was an outlier, with the second highest score recorded at 0.17 – much closer to the GZnivG range.

	Minimum	***Maximum***	***Range***	***Interquartile range***	***Mean***	***Standard deviation***
Total Experimental (n=130)	0.03	0.355	0.325	-	0.08	0.04
Gouzeaucourt (n=129)	0.03	0.22	0.18	0.04	0.09	0.03

Table 7. 35. Comparison of Edge straightness univariate statistics between Experimental and GZnivG assemblages.

L1/L (planform shape): The experimental assemblage's range and variance was greater than GZnivG's (see Table 7.36). In addition, given the contribution of points to the experimental sample, its mean was unsurprisingly lower and closer to Roe's (1968) threshold between pointed and ovate bifaces.

	Minimum	***Maximum***	***Range***	***Interquartile range***	***Mean***	***Standard deviation***
Total (n=130)	0.09	0.746	0.656	-	0.367	0.113
Gouzeaucourt (n=129)	0.22	0.64	0.42	0.08	0.43	0.07

Table 7. 36. Comparison of L1/L univariate statistics between Experimental and GZnivG assemblages.

7.4.1.2 Multivariate statistics

As shown in the experimental Discriminant Analysis, the accuracy with which one can predict skill-level was affected by handaxe shape. The quantitative criteria worked considerably better for ovates than for points (see Table 7.18). Therefore, because GZnivG was dominated by ovates, I used the 57.75% accuracy experimental ovate LDA loading data (including Intermediate Drafts) to analyse them (see Tables 7.19 and 7.20). I did not include Profile Asymmetry as it discriminated poorly between experimental skill-levels. Also, I omitted Flake Scar number but only because I was unable to collect the relevant GZnivG data due to Covid-19 restrictions.

Normalising the experimental LDA loadings was key to predicting the skill-level of the Acheulean knapper(s) at GZnivG. These were used to standardise every criterion score for individual bifaces on each axis and were added together to give a final combined axis score for every GZnivG biface. Then, based on the normalised ranges and means for experimental skill-levels on each axis, I generated preliminary thresholds for discriminating between Experts, Intermediates, and Novices. Due to the overlapping of quantitative scores between skill-levels adjacent to each other on the skill spectrum in my experimental investigation, I chose to include, in this case, borderline skill categories. This was predominantly driven by the quantitative similarity between Intermediates and Experts in my experimental sample. Additionally, there was intersection between

Intermediates and Novices. Therefore, boundary categories (Intermediate/Novice and Expert/Intermediate) straddled the ranges and means (normalised axis scores) of the original skill-levels (Novice, Intermediate, and Expert). In total, 5 skill-levels were assigned to GZnivG bifaces (see Table 7.38).

As Table 7.37 shows, Axis 1 explained most of the variance between the experimental bifaces and, thus, had the most discriminatory power. Though, to be more accurate in my initial assignments, where possible, I aimed to generate thresholds for each Axis and compared each Axis with another to help determine between skill-level classifications. However, as Table 7.38 illustrates, it was difficult to assign skill-level thresholds on Axes 2 and 3.

The chosen threshold for 'Expert' on Axis 1 was 0.14 (Table 7.38), which was the mean score for Experts on this Axis. However, as the mean score for Intermediates was not too dissimilar at 0.148, and the range was also very similar to Experts (see Tables 7.37 and 7.38), I reasoned that a score below the mean of 'Expert' corresponded with either Expert execution or leaps in Intermediate or Novice performance. As the experimental data showed, the majority of bifaces below 0.14 were Experts (53%), whilst only 31% of bifaces below 0.14 were Intermediate. I then imposed thresholds for the newly devised overlap category 'Expert/Intermediate' where the Axis 1 score corresponded to the higher half of the range that intersected with Experts and Intermediates and was above than the 'Expert' mean (0.1401 – 0.189). Intermediate classification was assigned to bifaces with Axis 1 scores between 0.18901 and 0.2. This was because this part of the Intermediate range did not overlap with the Expert Axis 1 range. In addition, I designated the Axis 1 threshold between Intermediate and Novice at 0.201 – 0.219. Although this range was

above the maximum Axis 1 score for Intermediate, it was within the range of both Intermediate Draft and Novice but below the mean for both. In fact, the ranges for Novice (0.118 – 0.316) and Intermediate Draft (0.168 – 0.341) coincided with Expert and Intermediate ranges but the combined number of bifaces for both skill-levels (Novice/Intermediate Draft) within the range of Expert or Intermediate in the experimental assemblage was 26%. Therefore, I reasoned that it was less likely that bifaces within the Expert and Intermediate ranges were made by Novices or were the result of a draft Intermediate biface attempt. Finally, the Axis 1 threshold for Novice was designated at or above the experimental Novice mean score of 0.22. Yet, the similarity in the ranges between Novice and Intermediate Draft meant that Axis 1 scores above 0.22 could denote either skill-level.

	Axis1 (73.72%)	**Axis2 (22.83%)**	**Axis3 (3.452%)**
Expert	Mean (0.14)	Mean (0.012)	Mean (-0.07)
Expert	Range (0.109 – 0.189)	Range (-0.02 – 0.049)	Range (-0.13 – 0.045)
Intermediate	Mean (0.148)	Mean (0.004)	Mean (-0.048)
Intermediate	Range (0.102 – 0.2)	Range (-0.01 – 0.02)	Range (-0.087 – 0.02)
Novice	Mean (0.22)	Mean (0.024)	Mean (-0.062)
Novice	Range (0.118 -0.316)	Range (-0.021 – 0.14)	Range (-0.173 – 0.018)
Intermediate Draft	Mean (0.23)	Mean (-0.012)	Mean (-0.066)

Interemdiate Draft	Range (0.168 – 0.341)	Range (-0.045 – 0.025)	Range -0.101 – 0.032)

Table 7. 37. Means and ranges of experimental skill-levels by normalised Axes scores.

For Axes 1 - 3, the thresholds were determined as so based on the above table:

	Axis1 (73.72%)	**Axis2 (22.83%)**	**Axis3 (3.45%)**
Expert	<0.14	n/a	n/a
Expert/Intermediate	0.1401 – 0.189	-0.02 – 0.012	n/a
Intermediate	0.18901 – 0.2	n/a	n/a
Intermediate/Novice	0.201 – 0.219	n/a	n/a
Novice	>0.22	<-0.02	n/a

Table 7. 38. Skill-level thresholds applied to GZnivG based on normalised experimental data.

With these thresholds I made a provisional assessment of the skill-level of GZnivG handaxes:

Skill-level	**Total**
Novice	19
Intermediate/Novice	26

Intermediate	22
Expert/Intermediate	55
Expert	7
Total	129

Table 7. 39. Initial quantitative skill-level assessment of GZnivG bifaces.

My initial classification suggested that the fewest represented skill-levels were of extreme Experts and extreme Novices, which partly supported Hypothesis 11 (see Chapter 6). The reasoning behind this postulate stemmed from the fact that because the true Novice stage is relatively short-lived and given that developed expertise takes considerable time and practice, most knappers in the archaeological record would likely be Intermediately skilled and somewhere along the spectrum to greater skill. Therefore, as a result, Experts and Novices would be more rarely represented than Intermediates.

By designating the most skilled threshold at below the mean of experimental Experts on Axis 1 (<0.14), and the least skilled threshold at above the mean of experimental Novices on Axis 1 (>0.22), I accounted for the highest and lowest (by my criteria) performing bifaces in the GZnivG assemblage. Between these lay the majority of biface Axis 1 scores that lay between the means of experimental Experts and Novices. These were partially, or fully, classified as Intermediate and existed along a spectrum from high-performing Intermediates (Expert/Intermediate) to Intermediates proper, to low-performing Intermediates (Intermediate/Novice). Naturally, this led to in more bifaces being partially classified as Intermediate than Novice or Expert, as per my rationale for Hypothesis 11. However, the similarity in Axis 1 scores between Experts and Intermediates meant there was a greater probability, when subjected to a LDA (see section 7.4.1.2.1), that several

Expert/Intermediates, or high-performing Intermediates, would be re-classified as Experts.

This outcome that 43% of GZnivG bifaces were initially classified as Expert/Intermediate (see Table 7.39) and, as such, fall within the Axis 1 range of Expert, is significant. It demonstrates, potentially counter to Hypothesis 11, that highly skilled knappers contributed most to the GZnivG assemblage. Conversely, this result is also in partial agreement with Hypothesis 11, in that it supports the status quo that no significant input was made by inexperienced knappers. Nevertheless, because the initial classification of GZnivG bifaces by their normalised Axis scores was based on thresholds chosen by me, they could hold inherent bias. It was necessary, therefore, to conduct a Discriminant Analysis to check my initial categorisations and improve the accuracy of them.

7.4.1.2.1 LDA

- 69.77% jackknifed classification accuracy.

Loadings

For each of the loading axes, the strongest discriminator corresponded to the normalised axes generated from experimental data.

	Axis 1 (96.16%)	**Axis 2 (3.14%)**	**Axis 3 (0.7%)**

Axis 1 (normalised)	48.152	-4.4908	-1.9949
Axis 2 (normalised)	6.1316	41.39	-24.681
Axis 3 (normalised)	4.0735	12.719	29.643

Table 7. 40. LDA 1 normalised GZnivG axis loading scores.

Confusion Matrix

	Novice	**Intermed/Novice**	**Intermed**	**Expert/Intermed**	**Expert**	**Pre-LDA Total**
Novice	12	7	-	-	-	19
Intermed/Novice	3	12	11	-	-	26
Intermediate	-	5	13	4		22
Expert/Intermed	-	-	7	46	2	55
Expert	-	-	-	-	7	7
Post-LDA Total	15	24	31	50	9	129

Table 7. 41. Confusion Matrix by normalised skill-level classification of GZnivG bifaces.

The post-LDA classification results in Table 7.41 illustrate that LDA helped to enhance the accuracy of the initial classifications presented above in Table 7.39. LDA highlighted and resolved my human error. For example, the LDA considered the Axis 2 and Axis 3 scores which, as Tables 7.37 and 7.38 showed, were almost impossible to provide skill

category thresholds for. This was why Axis 1 was predominantly used in my initial skill-level classification.

Comparison of the pre-LDA classifications and post-LDA classifications showed that, as predicted, Novice and Intermediate/Novice categories were reduced further, with more being included into the Intermediate category. This occurred along with some of the lower scoring Expert/Intermediate bifaces being reclassified as Intermediate too. Additionally, as anticipated, some higher scoring Expert/Intermediate bifaces were re-classified as Expert. However, this did not discount the fact that bifaces classified as Expert/Intermediate could equally likely have been made by either skill-level. Instead, the Expert category threshold identified the highest performing bifaces in the GZnivG assemblage.

Overall, this LDA was 69.77% accurate, but it did not tell the whole story. As can be seen in Table 7.41 above and Figure 7.22 below, reclassifications were not considerable and only occurred for skill-levels immediately adjacent to each other. There were no incorrect classifications that were considerably higher or lower than original classification. This suggested that, although only mostly using Axis 1 in my original skill-level classifications, my initial threshold placements were largely justified.

In sum, it was clear that LDA analysis improved my original skill-level classifications due to its enhanced ability to predict group membership through being able to incorporate Axes 2 and 3 scores into biface classification. However, the efficacy of the LDA classifications needed to be further evaluated. This is because they were based on data from an experimental assemblage of non-resharpened bifaces and, as previously

mentioned, GZnivG bifaces exhibited resharpening. Indeed, I proposed in Hypothesis 13 that smaller, more heavily re-sharpened GZnivG bifaces (McPherron, 1999) had the potential to be unfairly and negatively classified as being crafted by lower skill-levels due to their poorer refinement ratios. As such, the strength of the effect of resharpening on quantitative classification needed to be properly established by firstly conducting a post-classification univariate analysis (section 7.4.1.3) and, secondly, a comparative qualitative appraisal (section 7.4.2).

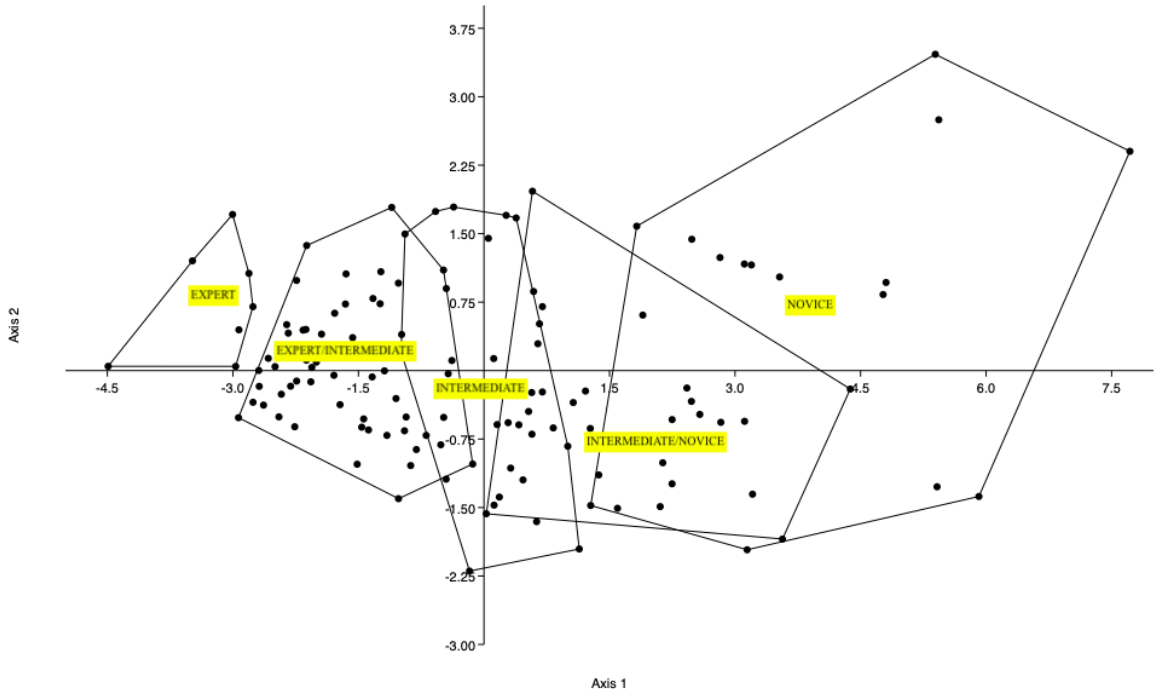


Figure 7. 24. Scatter plot of LDA re-classification of GZnivG bifaces.

7.4.1.3 Post-classification univariate statistical analysis

The post-LDA classification statistics for all quantitative criteria is detailed in Catalogue 3, Sheet 3.6. Examination of these criteria by LDA-assigned skill-level helped to evaluate the LDA classifications and any potential biases contributing to them.

Length:

As discussed above (see section 7.4.1.1), care must be taken when using data from experimental bifaces which did not undergo re-sharpening to classify GZnivG bifaces that did (McPherron, 1999, 1994; Soriano, 2000). The data showed that the Length of biface and/or extent of re-sharpening may have affected classification. Bifaces that exhibited less re-sharpening may have been classified as more skilful. Indeed, the mean length of GZnivG bifaces generally increased alongside skill-level category. In addition, the maximum length of an Intermediate/Novice biface was below the mean of those classified as Expert and Expert/Intermediate. Similarly, the shortest length of Experts was above the mean of Novice and Intermediate/Novice.

This outcome may be explained by Th/L ratio being the highest discriminating factor on Axis 1 in the main experimental LDA. However, there were shorter bifaces that were also classified in the highest skill-levels. The shortest Expert and Expert/Intermediate bifaces have similar lengths to the means of both Novice and Intermediate/Novice classified bifaces at GZnivG. Therefore, although length likely had some biasing effect on the GZnivG classifications, qualitative analysis will better illuminate the extent of the biasing impact of resharpening upon quantitative skill-level classifications at GZnivG.

Breadth:

The statistics indicated that, although the minimum breadths in the GZnivG assemblage belonged to bifaces classified as demonstrating less skill in their fabrication and mean scores showed that broader bifaces were more likely to be classified at Intermediate or above, all skill-levels possessed a good range of breadths. This potentially indicates that re-sharpening and/or the biface size was not overly biasing biface classification and other criteria (e.g., Edge Straightness and Planform Asymmetry) also played an important role.

Thickness:

Standard deviations for post-LDA thickness data demonstrated that, as in my experimental investigation (see Catalogue 4, Sheet 4.5, Table 3), higher skill-levels behaved more consistently than lower skill-levels. Indeed, Expert/Intermediate comprised the largest group by a considerable margin but displayed the second lowest variance.

Weight:

The post-LDA weight figures suggested weight did not affect skill-level classification. The range of Expert/Intermediate biface weights overlapped very closely with Novice classified bifaces, especially when the Expert/Intermediate maximum weight score outlier was removed. The minimum and maximum scores of these two skill-level categories were not too dissimilar either.

B/L ratio:

The post-LDA B/L ratio mean information supported the outcomes of the Length and Breadth data, suggesting that bifaces classified to lower skill-level categories (Novice and Intermediate/Novice) were, as per Gowlett (2011), on average shorter than those attributed to the higher skill-levels.

L1/L – Roe (1968) shape determination:

The very similar means for all skill-level categories in this variable intimated that producing a consistent final biface shape was a shared focus at GZnivG. Within this homogeneity, standard deviation scores indicated Novices acted more variably than other skill-level categories, potentially reflecting the lack of control typical of Novices that I discussed in Chapter 4. Alternatively, it could signify re-sharpening bias affecting biface classification.

Ratio of Refinement:

Generally, Ratio of refinement mean scores were lower (i.e., improved) among higher skill-level categories. This is an expected result of my quantitative methodology given that within the experimental dataset used to classify these GZnivG bifaces, the best

refinement scores were mostly produced by more experienced knappers. Furthermore, overlaps between skill-levels helped to reinforce my experimental observation that enhanced refinement scores can be achieved by lower-skilled individuals. Interestingly though, no intersection of refinement scores existed between Novice and Expert classified bifaces. This could signal resharpening bias in the results.

Thickness/Length ratio:

Like Refinement, Th/L ratio scores decreased (i.e., improved) as skill-level categories increased. Given that Th/L ratio was the most powerful discriminating factor in the experimental Axis 1 used to classify GZnivG bifaces, this was an expected result. It reflects my experiments which showed that the more refined handaxes were made by the more experienced experimental knappers.

Although some overlapping of the ranges of lower and higher skill-levels existed, this was mainly between skill-levels below that of Expert. Better TH/L ratio scores were achieved by lower-skill-levels, but less consistently, as the range and standard deviation results suggest. Again, this could imply resharpening bias in the results.

Planform Asymmetry:

The mean scores of all skill-levels were comparatively alike but were generally lower (i.e., improved) for superior skill-level categories. Surprisingly, however, the minimum

score for Experts was higher (i.e., worse) than for all other skill-levels. Moreover, the ranges and standard deviations for Expert and Expert/Intermediate classified bifaces were substantially smaller. Especially noteworthy is the low variance among bifaces classified as Expert/Intermediate and that constituted a large (39%) proportion of the GZnivG assemblage. Although further examination proved that Novice, Intermediate/Novice, and Intermediate skill-level categories all had upper limit outliers which increased their range and variance, bifaces classified as Expert or Expert/Intermediate possessed more constant symmetrical planform shapes.

Like the refinement ratios outlined above, these data clearly demonstrated the utility of a multivariate classification system. Reduced symmetry can be attributed to lower skill-levels as well as higher ones. However, when combined with scores from multiple criteria, anomalies in symmetry can be identified and overcome with regard to forming an accurate skill-level classification.

Profile Asymmetry:

Profile Asymmetry was not used to classify GZnivG bifaces due to its low discriminatory power among experimental bifaces. Despite being omitted in the LDA classification, the post-LDA data illustrated the reduced capacity of using experimental Profile Asymmetry data to distinguish between bifaces by the skill-level of their maker. The similarity of means and high degree of overlap in ranges between skill-level categories is particularly significant in this regard.

Edge Straightness:

Edge Straightness ranges, means, standard deviations, and minimum and maximum scores were all generally reduced among higher skill-level categories. Again, this was an expected result because straighter-edged experimental handaxes were generally made by more experienced knappers. Overall, Edge Straightness data revealed very little intersection between bifaces classified to higher and lower skill-level categories. This evidenced this criterion's strength in discriminating between bifaces by the skill-level of their maker.

7.4.1.4 Summary

Overall, post-LDA classification univariate analysis has shown that, in line with the experimental investigation, bifaces with lower Th/L ratio, Ratio of Refinement, Edge straightness, and Planform Asymmetry scores were generally classified as more highly skilled.

One of the main outcomes of this analysis, was that the prominence of TH/L as a discriminative criterion on the main LDA Axis 1 may cause larger bifaces to be categorised as more skilled. At GZnivG, resharpening could compound this effect, despite Length, Weight, and B/L ratio statistics demonstrating that multiple smaller bifaces were categorised among the higher skill-level categories. Consequently, it was

necessary to perform a qualitative assessment of the biases potentially inherent in the quantitatively derived GZnivG classifications.

7.4.2 Qualitative classification and comparison with quantitative results

As with the initial experimental investigation, I used the 3 broad skill-level categories of Expert, Intermediate, and Novice, to classify bifaces based on qualitative observations. These interpretations were recorded prior to quantitative analysis and detailed for each biface in Catalogue 1. Therefore, if I initially categorised a GZnivG biface as Intermediate, but the LDA results assigned an Intermediate/Novice or Expert/Intermediate classification, this was counted as correct. Consequently, what constituted an acceptable judgement could have varied considerably. Moreover, it is relevant to point out that the number of exact correlations between qualitative and quantitative classifications are likely to be affected by the inclusion of borderline skill-level categories (i.e., Expert/Intermediate and Intermediate/Novice) in my quantitative classification system. This is because qualitative assessment only used Expert, Intermediate, and Novice categories and, as such, each category cannot exactly correlate with a borderline classification.

7.4.2.1 GZnivG qualitative/quantitative comparison

The comparative accuracy of the two assessments was c. 90%. Qualitative description was necessary to verify whether a ‘correct’ quantitative assessment was only superficially

accurate or not. For example, a biface quantitatively classified as Expert/Intermediate could have correlated with a qualitative assessment of Intermediate, however, qualitative analysis could reveal mistakes indicative of lower-level Intermediate performance. Qualitative assessment thus provided clarity on the boundary ambiguity that existed in the quantitative classifications.

In total, only 13 bifaces had qualitative assessments that did not match with their quantitative classification (see Table 7.42). Whereas 45 GZnivG bifaces exhibited an exact match between assessments; this included all 9 bifaces quantitatively classified as Expert. Though most correct comparisons were with borderline quantitative classifications (n = 71), as explained above.

Below, I provide further analysis regarding each of these outcomes:

- A) Bifaces where quantitative and qualitative classifications did not correlate.
- B) Bifaces where quantitative and qualitative classifications correlated exactly.
- C) Ambiguous correlations between qualitative assessments and quantitative boundary classifications clarified by qualitative assessment.

It was unnecessary to present comparisons for every GZnivG handaxe here. The full list of correlations between quantitative and qualitative classifications is outlined in Table 7.42, which also shows the section (e.g., §1.011) of Catalogue 1 that corresponds to the biface description. The descriptions and qualitative observations in Catalogue 1 mirror the further detail provided in this section but for every GZnivG biface.

For Group A, I provide further analysis for all bifaces in that group (n = 13). This was to comprehensively cover the reasons for incorrect qualitative and quantitative comparisons. For Group B, I provide examples of bifaces that were both quantitatively and qualitatively categorised as Intermediate and Novice (n = 7), but there were too many to detail all of them here. I chose not to include further analysis regarding the 9 bifaces quantitatively classified as Expert because they were very clearly crafted with great skill (please see Table 7.42 and Catalogue 1). Most of the further detail provided is focused on Group C (n = 18), where the accuracy of borderline quantitative classifications is examined. My selections for all three Groups were designed to highlight the key strengths and weaknesses of my quantitative classification, to assess the impact of resharpening on classification accuracy, and to establish whether the high comparative accuracy (c. 90%) of quantitative and qualitative assessments reliably reflects what is going on at GZnivG.

Group A:

- GZ2802(15) (§1.011)
    - Qualitatively assessed as Novice but quantitatively assigned to the Expert/Intermediate category. It exhibited no clear indications of skilled work. Its quantitative score was the result of being made on flake. The thinness of the flake improved its Thickness/Length score, which was a predominant factor in the experimental data used to discriminate between skill-levels at GZnivG via LDA.
- GZ2812(22) (§1.033)
    - Qualitatively assessed as Novice but quantitatively assigned to the Intermediate category. It displayed no clear indications of proficient

craftspersonship and its quantitative score was also the result of being made on a flake as per GZ2802(15).

- GZ2812(26) (§1.037)
    - Qualitatively assessed as Expert but quantitatively assigned to the Intermediate category. Poor planform symmetry negatively affected the biface's quantitative classification. However, clear indications of skilled work were visible which validated my qualitative assessment.
- GZ2812(32) (§1.042)
    - Qualitatively assessed as Intermediate but quantitatively assigned to the Novice category. Quantitative classification showed the potential elongation bias inherent in the Thickness/Length ratio criterion. Consequently, very small bifaces like this one were occasionally negatively affected in terms of skill-level classification. Nevertheless, although edge maintenance was not Expert level, there were indications that the working of this very small piece was not by the hands of a Novice. The quantitative classification of Novice was unfair.
- GZ2812(36) (§1.045)
    - Qualitatively assessed as Expert but quantitatively assigned to Intermediate/Novice category. The biface's Th/L ratio, Asymmetry, and Edge Straightness negatively affected its quantitative classification. However, qualitative assessment showed the substantial expertise with which this previously endshocked biface had been recovered.
- GZ2832(6) (§1.051)
    - Qualitatively assessed as Intermediate but quantitatively assigned to the Novice category. I was unsure about whether to qualify this biface as

Intermediate or Novice, however, its small size and well-established plane of intersection influenced my decision towards Intermediate. Indications of skilled work were few though and the more quantitative conservative classification of Novice was not improbable.

- GZ2832(14) (§1.059)
    - Qualitatively assessed as Expert but quantitatively assigned to the Intermediate category. Qualitative assessment highlighted some inflexibility in the quantitative criteria. Due to a tranchet removal at the tip, which negatively affected the straightness of one edge, the classification was downgraded. After Th/L ratio, Edge straightness was the second most critical loading used to quantitatively classify GZnivG bifaces. Therefore, when affected it could lead to misclassification, even if it was affected by something that did not suggest poor craftspersonship.
- GZ2832(32) (§1.077)
    - Qualitatively assessed as Expert but quantitatively assigned to the Intermediate category. The biface's small size hindered its quantitative classification. However, its diminutive size, the quantitative scores it achieved, and the qualitative evidence for skilled manufacture, suggested its knapper possessed higher-than-Intermediate skill.
- GZ2832(33) (§1.078)
    - Qualitatively assessed as Expert but quantitatively assigned to the Intermediate category. This biface constituted a good example of how quantitative criteria cannot supply the level of context that qualitative assessment can provide. For instance, this biface was made on a small and relatively thick flake. This negatively affected refinement ratios. However,

as detailed in Catalogue 1, this knapper knew what to work and what to leave alone. They neatly worked the residual platform at the proximal end of the flake into a straight edge, which was not easy on such a small biface. In addition, they only minimally worked the distal end of the flake as it could already constitute a biface edge.

- GZ8015(10) (§1.087)
    - Qualitatively assessed as Expert but quantitatively assigned to the Intermediate category. Moderate Planform Asymmetry and Edge straightness scores were likely factors in the more conservative quantitative assessment of this biface. Qualitative assessment, on the other hand, showed skilled work on small material.
- GZ8015(12) (§1.089)
    - Qualitatively assessed as Expert but quantitatively assigned to the Intermediate category. The poor planform symmetry of this biface and the curvature of its edge (not zig-zagged sinuosity) were the principal factors affecting quantitative misclassification. Again, qualitative assessment illustrated the inflexibility of the quantitative criteria. A relatively poor Edge Straightness score was unfairly attributed to this biface. It did not occur due to poor edge management and pronounced sinuosity, but a general curvature of edges that were relatively well managed and well thinned.
- GZ8055(26) (§1.127)
    - Qualitatively assessed as Expert but quantitatively assigned to the Intermediate/Novice category. Like GZ8015(12), Edge straightness was a significant reason for misclassification. One edge exhibited considerable

sinuosity towards its butt, despite the edge being predominantly straight. Therefore, the sinuosity attributed to this biface via Edge straightness calculations did not represent continual sinuosity. Consequently, it should not be assessed as the result of unskilled performance. Again, this showed the occasional inflexibility of quantitative assessment without qualitative checks.

- GZ8055(28) (§1.129)
    - Qualitatively assessed as Intermediate but quantitatively assigned to the Novice category. Qualitative assessment highlighted the potential size bias involved in the quantitative assessment. It underlined how the very small size of the piece presented difficulties in reaching quantitative thresholds of higher skill-level categories. This was especially true when attempting to produce straight edges on small bifaces - there was little material to work with. Thus, Edge Straightness was the main contributing factor in the biface being marked down quantitatively. Qualitative appraisal stipulated the reasons why a quantitative classification of Novice was likely imbalanced given the knapper's ability to work with such small material.

B:

- GZ8055(4) (§1.107)
    - Qualitatively and quantitatively assessed as Intermediate. Despite good refinement ratios, which revealed that the knapper could thin effectively, the Edge Straightness score suggested the knapper was not in total control

of edge maintenance. Qualitative evaluation of unsuccessful thinning attempts, as well as missed thinning opportunities and sinuous edges, substantiated quantitative Edge Straightness findings. It thereby supported the methodological use of multiple criteria in a discriminant analysis to highlight simultaneously which actions a knapper could perform better than others.

- GZ8055(8) (§1.111)
    - Qualitatively and quantitatively assessed as Novice. Qualitative assessment validated quantitative scores which suggested the knapper was incapable of reliably maintaining the cutting edge and thinning the biface effectively.
- GZ8055(11) (§1.113)
    - Qualitatively and quantitatively judged to be made by a Novice. Quantitative scores were supported by qualitative assessment which established that the knapper was incapable of making a bifacial tool.
- GZ8055(20) (§1.121)
    - Qualitatively and quantitatively estimated to be a Novice product. Quantitative scores were supported by qualitative assessment which suggested the knapper could not create an even plane of intersection, thin the biface, or make straight edges.
- GZ2832(7) (§1.052)
    - Qualitatively and quantitatively considered to be crafted by an Intermediate. If this biface is compared to another qualitatively classified Intermediate biface such as GZ2832(6), it is clear that the Intermediate skill-level category straddled a significant spectrum and that two

qualitative assessments of the same skill-level classification can be based on disparate evaluations of skill (i.e., a higher and a lower level of Intermediate). Qualitatively, this knapper was skilled but not quite Expert. Quantitative scores agreed; however, this was mainly driven by the small size of the biface. If this biface was larger, I expect it would have been quantitatively re-classified as a boundary skill-level Expert/Intermediate.

- GZ2832(23) (§1.068)
    - Qualitatively and quantitatively assessed as Intermediate. Similar to GZ2832(7), if this biface was larger, I expect it would have been quantitatively re-classified as a boundary skill-level Expert/Intermediate.
- GZ2832(31) (§1.076)
    - Qualitatively and quantitatively attributed to the Intermediate skill-level category. It is an interesting biface given that its edge maintenance and straightness were good, but the thinning of the biface was inexpert. This was evidenced by numerous failed removals. Quantitative scores corroborated this qualitative explanation and highlighted the contradiction in performance between edge maintenance and thinning ability.

C (will be split into sections based on initial qualitative assessment):

i) Novice

- GZ2812(12) (§1.024)
    - Qualitatively assessed as Novice but quantitatively assigned to the Intermediate/Novice category. Its boundary quantitative classification was

driven by Refinement scores. However, qualitative assessment highlighted multiple instances of unskilled performance. In fact, my qualitative interpretation suggested this biface was the product of a child playing with a small piece of flint.

ii) Intermediate

- GZ2802(7) (§1.004)
    - Qualitatively assessed as Intermediate but quantitatively assigned to the Expert/Intermediate category. Qualitative and diacritical analysis demonstrated that there was limited evidence of Expert performance or a coherent flaking plan and, as such, the boundary quantitative classification should come down on side of Intermediate.
- GZ2802(13) (§1.009)
    - Qualitatively assessed as Intermediate but quantitatively appointed to the Intermediate/Novice category. Qualitative assessment identified that post-depositional factors affected quantitative criteria and thus suggests the quantitative boundary classification was perhaps unfair.
- GZ2802(17) (§1.013)
    - Qualitatively assessed as Intermediate but quantitatively allocated to the Intermediate/Novice category. Its small size and 'S' shaped edges negatively affected quantitative classification. Qualitative assessment insinuated this biface was more appropriately classified as an Intermediate or Expert/Intermediate, given the reduced dimensions of the raw material being worked, and the thinness and edge straightness achieved.

- GZ2812(3) (§1.016)
    - Qualitatively assessed as Intermediate but quantitatively assigned to the Intermediate/Novice category. Qualitative assessment exposed multiple instances of mistakes and poor performance. Therefore, a boundary quantitative skill-level categorisation for this biface was reasonable.
- GZ2812(7) (§1.019)
    - Qualitatively assessed as Intermediate but quantitatively calculated to belong to the Intermediate/Novice category. Qualitative assessment revealed that this biface's quantitative boundary classification was being affected predominantly by the size of the biface. There was occasional evidence of inexpert action, but not Novice performance.
- GZ2812(16) (§1.028)
    - Qualitatively assessed as Intermediate but quantitatively assigned to the Intermediate/Novice category. Qualitative assessment again revealed the occasional rigidity of the quantitative criteria. Although the biface's Edge Straightness score was not good, and the thinning could have been improved, the knapper knew what they wanted to achieve and could do it but not proficiently. This was demonstrated by the attempted removal of a tranchet. Overall, the behaviours exhibited by this biface were indicative of above Novice performance.
- GZ2812(31) (§1.041)
    - Qualitatively assessed as Intermediate but quantitatively ascribed to the Intermediate/Novice category. Quantitative classification did not realistically represent the skill displayed by the knapper when working this

small biface. Again, resharpening/size bias had a negative effect on the quantitative classification.

- GZ2832(13) (§1.058)
    - Qualitatively assessed as Intermediate but quantitatively assigned to the Intermediate/Novice category. The knapper's capacity to thin properly was influenced by the diminutive size of the biface and chalky inclusions present in the material. Both negatively affected Refinement ratios and qualitative assessment suggested it was reasonably well made and more Intermediate than Novice.
- GZ2832(20) (§1.065)
    - Qualitatively assessed as Intermediate but quantitatively attributed to the Expert/Intermediate category. The knapper found it difficult to thin one face, creating a plano-convex form that was atypical at GZnivG. This was the reason I qualitatively evaluated it as Intermediate-level, but, overall, the knapper performed quite skilfully, thus vindicating its quantitative classification.
- GZ8015(2) (§1.080)
    - Qualitatively assessed as Intermediate but quantitatively assigned to the Expert/Intermediate category. This biface was generally well made but the knapper did not quite display Expert performance. Hence, a boundary quantitative skill-level classification was reasonable. Nonetheless, this was one of the larger bifaces in the sample and other smaller bifaces, such as GZ2802(17), demonstrated comparably skilled behaviour but were not classified similarly. This biface was another that exposed the length bias

inherent in the quantitative criteria used as part of the discriminant analysis.

- GZ8015(6) (§1.084)
    - Qualitatively assessed as Intermediate but quantitatively allotted to the Expert/Intermediate category. Like GZ8015(2), qualitative analysis revealed quantitative classification was acceptable, despite there also being evidence of biface length biasing classification.
- GZ8015(8) (§1.086)
    - Qualitatively assessed as Intermediate but quantitatively assigned to the Intermediate/Novice category. This biface constituted an example where qualitative assessment did not match quantitative classification, notwithstanding the borderline correct classification. Quantitative classification was driven by the small size of the biface, however, qualitative assessment shows the knapper was, although not Expert, certainly not a Novice or having demonstrated inexperienced, or poor, performance.
- GZ8015(21) (§1.098)
    - Qualitatively assessed as Intermediate but quantitatively consigned to the Intermediate/Novice category. The very small size of this biface affected its quantitative classification; however, quantitative criteria did not categorise it as fully Novice. This was important in demonstrating that although TH/L ratio did have a significant impact on quantitative classification, the other criteria in the discriminant analysis also contributed importantly to finding a skill-level grouping more in line with

the qualitative assessment. Indeed, qualitative assessment shows that this biface was more reflective of Intermediate than Novice performance.

- GZ8015(23) (§1.100)
    - Qualitatively assessed as Intermediate but quantitatively assigned to the Expert/Intermediate category. Qualitative identification of inexpert performance suggested that the biface's quantitative boundary classification may have been driven by it being made on a flake, which explains the good Th/L score. Consequently, quantitative classifications clearly sometimes require context provided by qualitative assessment. Moreover, bifaces made on flakes may occasionally be unfairly quantitatively evaluated as being made by more highly skilled knappers due to their improved Refinement ratios. Examples such as GZ2802(15) and GZ2812(22), among others, also attested to this.

iii) Expert

- GZ2802(9) (§1.006)
    - Qualitatively assessed as Expert but quantitatively appointed to the Expert/Intermediate category. Once more, the inflexibility of quantitative criteria was responsible for the quantitative marking-down of this biface. Quantitative classification was adversely affected by poor planform symmetry, but asymmetry was predominantly located at the butt only. Therefore, as qualitative assessment intimated, this biface was more reasonably considered to be the product of a craftsperson possessing an Expert knapping level. Although the discrepancy between quantitative and qualitative classifications was not overly marked for this biface, the cause

of the discordance brings into question the utility of using planform symmetry in quantitative assessment of skill-levels moving forward.

- GZ2802(14) (§1.010)
    - Qualitatively assessed as Expert but quantitatively assigned to the Expert/Intermediate category. Qualitative assessment shows that little to no mistakes were made by the knapper. Thus, its boundary classification was again being driven by length bias inherent in the normalised experimental LDA loadings used to analyse GZnivG.
- GZ2832(25) (§1.070)
    - Qualitatively assessed as Expert but quantitatively allocated to the Expert/Intermediate category. Quantitative classification did not capture the qualitatively observable skill with which the knapper has worked a very difficult material. Indeed, the knapper operated around an inclusion/void towards the tip to create a very skilled handaxe. The medium size of the handaxe, and asymmetry at the tip may also have been key contributing factors behind this boundary classification.

One of the main outcomes of this quantitative/qualitative comparison was that biases within the resulting quantitative classifications were detected via qualitative assessment. Firstly, when bifaces were made on a flake, this could have the effect (e.g., GZ2802(15) and GZ2812(22)) of improving the refinement ratios beyond the level commensurate with the qualitative level of behaviour the biface exhibited. Consequently, method of manufacture (i.e., reduction of a whole nodule, or production on a flake) influenced quantitative classification. Additionally, qualitative analysis showed that, as a result of

resharpening, very small bifaces, below c. 70mm in length at GZnivG (e.g., GZ8055(28) and GZ2812(32)), were often negatively classified by quantitative criteria. Consequently, although my experimental data showed Novices to make smaller and thicker handaxes, the smaller and less refined GZnivG bifaces were not always necessarily produced by lower skill-levels.

Despite these biases, most of the comparative classifications are accurate, as Table 7.42 shows. This demonstrates the utility of employing a mixed methods approach. Small and large bifaces were correctly classified for all skill-level categories. It was only very small, heavily resharpened GZnivG bifaces that possessed a greater chance of being negatively misclassified. Furthermore, as section 'A' above illustrated, other criteria (Edge straightness and Planform Asymmetry) were also revealed, by the context qualitative assessment provided, to be responsible for misclassifications.

Overall, qualitative analysis demonstrated that the efficacy of the quantitative criteria used to discriminate between bifaces based on the maker's skill-level could potentially be improved if applied to a non-resharpened archaeological assemblage. In Chapter 8, to test this idea I use a selection of Boxgrove bifaces where resharpening was not employed. However, before this I conduct a secondary qualitative assessment of the diacritical flaking patterns exhibited by the GZnivG bifaces (see below). Using the qualitative classifications outlined in Table 7.42, I evaluate whether any differences in the way bifaces were made can be evinced from diacritical analysis.

Biface no.	Qualitative classification	Quantitative classification (Jackknifed)	Catalogue 1 reference	Correct/Incorrect
GZ2802(4)	Expert	Expert/Intermediate	§1.001	✓
GZ2802(5)	Intermediate	Expert/Intermediate	§1.002	✓
GZ2802(6)	Intermediate	Intermediate	§1.003	✓
GZ2802(7)	Intermediate	Expert/Intermediate	§1.004	✓
GZ2802(8)	Intermediate	Expert/Intermediate	§1.005	✓
GZ2802(9)	Expert	Expert/Intermediate	§1.006	✓
GZ2802(11)	Intermediate	Intermediate	§1.007	✓
GZ2802(12)	Intermediate	Intermediate	§1.008	✓
GZ2802(13)	Intermediate	Intermediate/Novice	§1.009	✓
GZ2802(14)	Expert	Expert/Intermediate	§1.010	✓
GZ2802(15)	Novice	Expert/Intermediate	§1.011	×
GZ2802(16)	Novice	Novice	§1.012	✓
GZ2802(17)	Intermediate	Intermediate/Novice	§1.013	✓
GZ2812(1)	Expert	Expert/Intermediate	§1.014	✓
GZ2812(2)	Expert	Expert/Intermediate	§1.015	✓
GZ2812(3)	Intermediate	Intermediate/Novice	§1.016	✓
GZ2812(4)	Intermediate	Intermediate	§1.017	✓
GZ2812(6)	Expert	Expert	§1.018	✓
GZ2812(7)	Intermediate	Intermediate/Novice	§1.019	✓
GZ2812(8)	Intermediate	Intermediate	§1.020	✓
GZ2812(9)	Expert	Expert	§1.021	✓
GZ2812(10)	Intermediate	Intermediate	§1.022	✓
GZ2812(11)	Intermediate	Intermediate/Novice	§1.023	✓
GZ2812(12)	Novice	Intermediate/Novice	§1.024	✓
GZ2812(13)	Intermediate	Intermediate/Novice	§1.025	✓
GZ2812(14)	Expert	Expert	§1.026	✓

GZ2812(15)	Novice	Intermediate/Novice	§1.027	✓
GZ2812(16)	Intermediate	Intermediate/Novice	§1.028	✓
GZ2812(18)	Intermediate	Intermediate	§1.029	✓
GZ2812(19)	Expert	Expert	§1.030	✓
GZ2812(20)	Expert	Expert/Intermediate	§1.031	✓
GZ2812(21)	Novice	Novice	§1.032	✓
GZ2812(22)	Novice	Intermediate	§1.033	×
GZ2812(23)	Expert	Expert/Intermediate	§1.034	✓
GZ2812(24)	Expert	Expert	§1.035	✓
GZ2812(25)	Expert	Expert	§1.036	✓
GZ2812(26)	Expert	Intermediate	§1.037	×
GZ2812(27)	Intermediate	Intermediate	§1.038	✓
GZ2812(28)	Novice	Novice	§1.039	✓
GZ2812(29)	Intermediate	Intermediate	§1.040	✓
GZ2812(31)	Intermediate	Intermediate/Novice	§1.041	✓
GZ2812(32)	Intermediate	Novice	§1.042	×
GZ2812(34)	Intermediate	Intermediate	§1.043	✓
GZ2812(35)	Intermediate	Intermediate/Novice	§1.044	✓
GZ2812(36)	Expert	Intermediate/Novice	§1.045	×
GZ2832(1)	Expert	Expert/Intermediate	§1.046	✓
GZ2832(2)	Expert	Expert/Intermediate	§1.047	✓
GZ2832(3)	Expert	Expert/Intermediate	§1.048	✓
GZ2832(4)	Expert	Expert/Intermediate	§1.049	✓
GZ2832(5)	Expert	Expert/Intermediate	§1.050	✓
GZ2832(6)	Intermediate	Novice	§1.051	×
GZ2832(7)	Intermediate	Intermediate	§1.052	✓
GZ2832(8)	Expert	Expert/Intermediate	§1.053	✓
GZ2832(9)	Expert	Expert	§1.054	✓
GZ2832(10)	Expert	Expert/Intermediate	§1.055	✓

GZ2832(11)	Intermediate	Expert/Intermediate	§1.056	✓
GZ2832(12)	Expert	Expert	§1.057	✓
GZ2832(13)	Intermediate	Intermediate/Novice	§1.058	✓
GZ2832(14)	Expert	Intermediate	§1.059	×
GZ2832(15)	Intermediate	Expert/Intermediate	§1.060	✓
GZ2832(16)	Intermediate	Intermediate	§1.061	✓
GZ2832(17)	Novice	Novice	§1.062	✓
GZ2832(18)	Expert	Expert/Intermediate	§1.063	✓
GZ2832(19)	Intermediate	Expert/Intermediate	§1.064	✓
GZ2832(20)	Intermediate	Expert/Intermediate	§1.065	✓
GZ2832(21)	Intermediate	Expert/Intermediate	§1.066	✓
GZ2832(22)	Intermediate	Intermediate/Novice	§1.067	✓
GZ2832(23)	Intermediate	Intermediate	§1.068	✓
GZ2832(24)	Intermediate	Expert/Intermediate	§1.069	✓
GZ2832(25)	Expert	Expert/Intermediate	§1.070	✓
GZ2832(26)	Intermediate	Intermediate/Novice	§1.071	✓
GZ2832(27)	Expert	Expert/Intermediate	§1.072	✓
GZ2832(28)	Intermediate	Intermediate	§1.073	✓
GZ2832(29)	Expert	Expert/Intermediate	§1.074	✓
GZ2832(30)	Intermediate	Intermediate	§1.075	✓
GZ2832(31)	Intermediate	Intermediate	§1.076	✓
GZ2832(32)	Expert	Intermediate	§1.077	×
GZ2832(33)	Expert	Intermediate	§1.078	×
GZ8015(1)	Expert	Expert/Intermediate	§1.079	✓
GZ8015(2)	Intermediate	Expert/Intermediate	§1.080	✓
GZ8015(3)	Expert	Expert/Intermediate	§1.081	✓
GZ8015(4)	Intermediate	Intermediate	§1.082	✓
GZ8015(5)	Novice	Novice	§1.083	✓
GZ8015(6)	Intermediate	Expert/Intermediate	§1.084	✓

GZ8015(7)	Intermediate	Intermediate	§1.085	✓
GZ8015(8)	Intermediate	Intermediate/Novice	§1.086	✓
GZ8015(10)	Expert	Intermediate	§1.087	×
GZ8015(11)	Intermediate	Expert/Intermediate	§1.088	✓
GZ8015(12)	Expert	Intermediate	§1.089	×
GZ8015(13)	Expert	Expert/Intermediate	§1.090	✓
GZ8015(14)	Expert	Expert/Intermediate	§1.091	✓
GZ8015(15)	Intermediate	Intermediate	§1.092	✓
GZ8015(16)	Expert	Expert/Intermediate	§1.093	✓
GZ8015(17)	Expert	Expert/Intermediate	§1.094	✓
GZ8015(18)	Expert	Expert/Intermediate	§1.095	✓
GZ8015(19)	Expert	Expert/Intermediate	§1.096	✓
GZ8015(20)	Expert	Expert/Intermediate	§1.097	✓
GZ8015(21)	Intermediate	Intermediate/Novice	§1.098	✓
GZ8015(22)	Intermediate	Intermediate/Novice	§1.099	✓
GZ8015(23)	Intermediate	Expert/Intermediate	§1.100	✓
GZ8015(24)	Intermediate	Expert/Intermediate	§1.101	✓
GZ8015(25)	Novice	Intermediate/Novice	§1.102	✓
GZ8015(26)	Expert	Expert/Intermediate	§1.103	✓
GZ8055(1)	Expert	Expert/Intermediate	§1.104	✓
GZ8055(2)	Expert	Expert/Intermediate	§1.105	✓
GZ8055(3)	Expert	Expert/Intermediate	§1.106	✓
GZ8055(4)	Intermediate	Intermediate	§1.107	✓
GZ8055(5)	Expert	Expert/Intermediate	§1.108	✓
GZ8055(6)	Expert	Expert/Intermediate	§1.109	✓
GZ8055(7)	Intermediate	Intermediate/Novice	§1.110	✓
GZ8055(8)	Novice	Novice	§1.111	✓
GZ8055(9)	Intermediate	Intermediate/Novice	§1.112	✓
GZ8055(11)	Novice	Novice	§1.113	✓

GZ8055(12)	Novice	Novice	§1.114	✓
GZ8055(13)	Expert	Expert/Intermediate	§1.115	✓
GZ8055(14)	Expert	Expert/Intermediate	§1.116	✓
GZ8055(15)	Intermediate	Intermediate	§1.117	✓
GZ8055(16)	Intermediate	Intermediate	§1.118	✓
GZ8055(17)	Intermediate	Intermediate	§1.119	✓
GZ8055(19)	Expert	Expert	§1.120	✓
GZ8055(20)	Novice	Novice	§1.121	✓
GZ8055(21)	Intermediate	Intermediate	§1.122	✓
GZ8055(22)	Intermediate	Intermediate/Novice	§1.123	✓
GZ8055(23)	Novice	Novice	§1.124	✓
GZ8055(24)	Novice	Novice	§1.125	✓
GZ8055(25)	Intermediate	Intermediate/Novice	§1.126	✓
GZ8055(26)	Expert	Intermediate/Novice	§1.127	×
GZ8055(27)	Novice	Novice	§1.128	✓
GZ8055(28)	Intermediate	Novice	§1.129	×

Table 7. 42. GZnivG: comparison of quantitative and qualitative skill-level assessments.

7.4.2.2 Diacritical analysis

Diacritical analysis transformed qualitative observation into quantitative data (see Table 7.43). As per Lamotte's (2012, 2011) method, the sequence of removals can be described in 8 ways (see Figure 6.5). Some of these 8 categories can be joined together (i.e., there are multiple Bilateral and Trilateral removal sequence categories). I assessed each biface and gave a diacritical assessment of both faces (see Catalogue 3, Sheet 3.4). The combinations of both faces can be distilled into the categories outlined in Table 7.43.

	Qualitative skill-level classification				
	Novice	Intermediate	Expert	Total	%
Unilateral and Unilateral	1	0	0	1	1%
Unilateral and Bilateral	0	5	2	7	5%
Unilateral and Trilateral	0	1	0	1	1%
Unilateral and Centripetal	0	1	0	1	1%
Unilateral and Diverse	3	1	0	4	3%
Bilateral and Bilateral	2	13	21	36	28%
Bilateral and Trilateral	0	4	4	8	6%
Bilateral and Centripetal	2	4	3	9	7%
Bilateral and Diverse	1	1	1	3	2%
Trilateral and Trilateral	1	4	2	7	5%
Trilateral and Centripetal	0	3	1	4	3%
Trilateral and Diverse	1	1	0	2	2%
Centripetal and Centripetal	2	10	14	26	20%
Centripetal and Diverse	0	2	0	2	2%
Diverse and Diverse	4	10	4	18	14%
Total	17	60	52	129	

Table 7. 43. Diacritical assessment by qualitative skill-level classification.

Table 7.43 shows that two diacritical categories predominated among the GZnivG assemblage (Bilateral and Bilateral, and Centripetal and Centripetal). These two

groupings (out of 15) were found on 48% of GZnivG bifaces. Among Expert and Intermediate qualitatively classified bifaces, these two were the first and second (joint second for Intermediate) most common diacritical patterns. However, among Novice classified bifaces these configurations appear less frequently (joint third most common categories). For Novices, a more randomised removal approach (Diverse and Diverse category) was most common, with the Unilateral and Diverse category being the second most applied diacritical pattern.

Although the 'Diverse and Diverse' category was the other joint second most common applied configuration for Intermediate bifaces, this approach was infrequently visible among Expert classified bifaces. In addition, whereas bifaces with partially Unilateral diacritical arrangements accounted for 24% of the Novice assemblage, they only accounted for 13% of the Intermediate assemblage, and 4% of the Expert assemblage. Moreover, differences between the skill-levels can be seen in the number of bifaces that employed Bilateral and Centripetal patterns on at least one face. At least one bilaterally worked face was found on 60% of Expert handaxes, 45% of Intermediate handaxes, and on only 29% of Novice handaxes. Additionally, 35% of Expert and 33% of Intermediate bifaces yielded at least one centripetally worked face, whereas Novice bifaces that exhibited at least one centripetally worked face accounted for only 24% of their assemblage.

The greater variation in removal patterns among Novice bifaces suggested, in line with experimental and archaeological quantitative data presented above, that Novices knapped with less control over their performance. In Table 6, Sheet 3.6, Catalogue 3, we see that the shapes of bifaces classified as Novice at GZnivG were more varied than higher skill-

levels, just as their patterns of flake removal were. Therefore, shape and diacritical data split by skill-level classification, implied that Novices found it more difficult to adhere to shared constraints on handaxe manufacture that existed at GZnivG. The greater homogeneity of shape and removal patterns among bifaces classified as Expert and Intermediate suggested that higher skill-levels were able to constrain their emerging intentions more successfully in relation to a common goal. Certainly, the results of the Blind test in section 7.3 would support this interpretation.

7.5 Summary

Overall, the above results have shown it was possible to discriminate between bifaces based on the skill-level of their maker through a mixed method approach. However, the results were not always reliable. Quantitative analysis of experimental bifaces has generated data that can be used to confidently distinguish between archaeological bifaces whose makers were Expert/Intermediate or were Novice. However, there was a clear overlap between Expert and Intermediate level knappers which reduced the confidence of assigning skill classifications to these two categories using solely quantitative data. Therefore, qualitative analysis was still necessary to provide nuance and context to the assessment of a biface by illustrating examples of skilled behaviour that were impossible to capture with the quantitative criteria. Therefore, further research is required to discover more effective quantitative ways to distinguish between bifaces more accurately.

Additionally, I have shown through experiments that both practice and tuition contributed to occasional ephemeral improvements in performance by less experienced knappers. This resulted in quantitative, and sometimes qualitative (e.g., biface CS_LIV2.13_made

by an Intermediate) examples of parity with Expert knappers. This added to uncertainty surrounding correctly classifying the experience-level of the knapper based solely on the biface without any other contextual information such as the by-products of manufacture, something that, as discussed above, is a commonplace occurrence in the Acheulean.

Furthermore, I have demonstrated that biface shape affected the accuracy of my quantitative criteria. I have shown that the skill criteria used here more accurately categorised ovate bifaces than pointed ones. This important finding suggests that, regarding my quantitative criteria, biface morphology affected what was achievable by knappers, especially in terms of refinement and symmetry. Consequently, ovates and pointed bifaces will yield different quantitative thresholds for separating skill-levels. Moreover, my analysis of the archaeological assemblage from GZnivG suggested that among resharpened bifaces, some of the quantitative criteria in my analysis were less diagnostic in discriminating between bifaces below a certain size threshold based on skill-level of their maker.

These issues will be discussed in greater detail in Chapter 8, where I assess the significance of the results, the usefulness of the methodology for future research, and suggest how my approach could be improved. Before concluding this chapter, however, I will briefly refer to the hypotheses developed for the Experimental and Archaeological studies (see Chapter 6) and discuss whether to confirm or reject them on the basis of my results.

7.5.1 Main experimental investigation

My first experimental hypothesis (H^1) can only be partly corroborated because the results show that although Novices could often be accurately distinguished from Expert or Intermediate handaxes, separating Expert and Intermediate handaxes was more difficult.

As a consequence, Hypothesis 2 (H^2) can be rejected. Counter to my expectations, most quantitative misclassifications were between Expert and Intermediate knapped bifaces. Finding a solution to separating these two skill-levels more reliably using quantitative criteria is an important avenue for further investigation.

Hypothesis 3 (H^3) can be accepted. Both univariate and multivariate statistics outlined above (sections 7.2.1.1 and 7.2.1.3) suggested that even though the gap with Intermediates was sometimes small, Experts exhibited the least variable performance against a number of criteria.

For the Blind Test, Hypothesis 4 (H^4), can also be accepted. My quantitative criteria clearly distinguished between the two experimental skill-levels. Although the sample size was small, it is significant that replication of a template helped to illustrate the differences between the two Blind Test knappers more than the original experimental investigation did. The import of this outcome for future research will be discussed further in the next and final chapter.

Hypothesis 5 (H^5) can be accepted too. As Table 7.18 shows, for all LDA combinations pointed bifaces were less accurately predicted using my quantitative criteria than were ovate bifaces.

Hypotheses 6 and 7 (H^6 & H^7) can also be confirmed given that the Novice (Experimental knapper 7) quantitative scores (sections 7.2.5.1 and 7.2.5.2) improved with practice over time, as well as under direct instruction.

Improvement over time with practice was also shown via the quantitative scores achieved by the Intermediate (Experimental knapper 4). Therefore, Hypothesis 8 (H^8) can be partly accepted because bifaces that the Intermediate made in their third year of practice could be distinguished from handaxes made during their earlier knapping years. I mention 'partly' accepted because, although the general trend was improvement by this knapper, 33% of bifaces made by the Intermediate in their first year were misclassified as handaxes made in their third year (after Expert instruction). This may suggest potential leaps and dips in performance as per Gray and Lindstedt (2017).

However, Hypothesis 9 (H9) can be fully accepted, with the biface that Experimental knapper 4 made under instruction exhibiting clear quantitative and qualitative improvement in performance.

7.5.2 Archaeological Hypotheses

Regarding my hypotheses for the assemblage from GZnivG, Hypothesis 10 (H^{10}) can be accepted. Multiple skill levels could be detected based on the experimental data, despite the effect of underlying biases on classification.

Hypothesis 11 (H^{11}) can also be confirmed. If we take the thresholds outlined in section 7.4.1.2, only 24 of 129 GZnivG bifaces fitted the extreme Novice or Expert categories. Conversely, there was significant overlap between Expert and Intermediate. Indeed, qualitative analysis was equally equivocal in distinguishing Experts and Intermediates, despite showing Novices to be rarer than clear Experts. If we use the quantitative thresholds as an index of skill spectrum, then there were relatively fewer extremely highly skilled (<0.14 on Axis 1) and low skilled (>0.22 on Axis 1) bifaces at GZnivG. The relative under-representation of Novices in both quantitative and qualitative analysis could mean that reaching borderline Intermediate quantitative scores was a temporally short journey, or that they were genuinely in the minority of GZnivG knappers.

Hypothesis 12 (H^{12}) can largely be accepted too. As Table 7.42 shows, there is a c. 90% correlative accuracy between qualitative and quantitative skill-level classifications of GZnivG bifaces. However, that meant in approximately 10% of cases the qualitative and quantitative classifications did not agree. Moreover, there were examples where quantitative boundary classifications (i.e., Expert/Intermediate and Intermediate/Novice) suggested the knapper was more skilled than what the qualitative assessment intimated, and vice versa.

As discussed above, the elongation/resharpening bias outlined in Hypothesis 13 (H^{13}) can also be confirmed to exist at GZnivG. LDA was less effective in quantitatively classifying smaller, more resharpened GZnivG bifaces correctly.

Finally, diacritical analysis showed that variation of flaking patterns was not markedly variable at GZnivG, but most of the diacritical variability in the assemblage was generated by Novices – thus corroborating Hypothesis 14 (H^{14}).

8. Chapter 8 – Discussion and Interpretation

8.1 Introduction

The aim of this thesis has been to propose a method for discriminating between archaeological bifaces by the perceived skill-level they exhibit. Key to this was examining an archaeological assemblage where a consistent, shared approach to biface manufacture could be demonstrated. I have reasoned that skill is something that is best understood as relational between associated practitioners in a group that shares a common pursuit.

Hence, I chose to examine bifaces from GZnivG. Previous studies of bifaces from the site (Field, 2005; McPherron, 1999; Soriano, 2000) have shown they were very similarly manufactured and shaped. But I have gone further by demonstrating that planform shape was consistently applied regardless of biface size or resharpening intensity. Moreover, my diacritical results (see section 7.4.2.2) supported Soriano's (2000) claim that the Gouzeaucourt knappers performed similar techno-functional patterns of behaviour. Therefore, GZnivG was deemed a highly suitable assemblage to investigate skill-level differences among Acheulean knappers.

In order to interpret the GZnivG data and discriminate between bifaces by the skill-level of their makers, it was necessary to use data from modern knappers, where varying abilities and experience levels are known. My mixed methods approach, using quantitative and qualitative analyses to classify bifaces, successfully discriminated between higher and lower skill-levels among my experimental sample. Misclassifications

inevitably occurred, but mainly involved adjacent skill categories. For example, it was very rare for Novice and Expert handaxes to be misclassified as each other, but both, especially Experts, overlapped with Intermediates more regularly.

Importantly, when using my quantitative criteria to assess skill-level from bifaces alone, my main experimental results identified fundamental problems associated with accurately identifying skill-level thresholds along the skill spectrum. Correct classifications were susceptible to being influenced by dips/inconsistencies in performance, as well as considerable improvements via direct, step-by-step Expert instruction. Moreover, the lack of precise constraints imposed upon knappers in my original experiment highlighted the effect that biface shape/morphology had on the effectiveness of my quantitative criteria to discriminate between bifaces by skill-level. This was a significant finding given that I, and others (Shipton, 2010; Shipton and White, 2020) have consistently argued that normative standards of handaxe manufacture were shared among knappers in the Acheulean. Indeed, my Blind Test indicated that applying constraints to manufacture may improve the effectiveness of my criteria to quantitatively distinguish between Expert and Intermediate. It was therefore crucial to gather these observations before applying my quantitative criteria to analyse GZnivG bifaces.

A variable that I could not allow for experimentally, however, was the prevalence of resharpening present among the GZnivG assemblage. Qualitative analysis supported my prediction that resharpening would influence quantitative classification, particularly among smaller, more heavily resharpened archaeological bifaces. This important finding suggested that my quantitative criteria would work more effectively on an assemblage where re-sharpening was less prominent and would thus be less reliant on qualitative

analysis to provide clarity on classification. In order to test this idea, I therefore extended my study to include an assemblage from Boxgrove (see section 8.3).

Once the key findings have been discussed in more detail in the next section, and my methodology is tested on a predominantly non-resharpened assemblage in section 8.3, I will go on to outline future research directions that grew from this project. I conclude this thesis by commenting on what my assessment of the distribution of skill-levels at GZnivG indicates regarding the use of the site.

8.2 Key findings and limitations

As Table 7.18 shows, I tested numerous combinations of quantitative criteria on the entire experimental assemblage, and by biface shape (ovate or pointed). A key discovery was that my criteria were better at accurately classifying ovates than pointed bifaces by skill-level category. This was a significant finding given that GZnivG bifaces were ovate.

The criteria I used in LDA (Linear Discriminant Analysis) 1 to analyse experimental ovate bifaces gave an accuracy score of 58% (see Table 7.41). Although this result can clearly be improved, it is a noteworthy result. Indeed, on closer inspection (see Tables 7.43 and 7.44) the significance of this result became even more apparent; misclassification of bifaces mostly only occurred between skill-levels adjacent to each other. In fact, there were only very few exceptions to this pattern. I found that Expert bifaces were never misclassified as Novice, Intermediate and Intermediate Draft handaxes were notably over 70% accurately classified, and Novice bifaces were

predominantly misclassified to the bordering Intermediate category. In total, the key outcome was that only 6 out of 71 ovate experimental bifaces were misclassified beyond adjacent skill-level boundaries. This result indicated the quantitative criteria used here were, in the main, effective at discriminating between knappers of different levels of experience. Following expectation, lower skill-levels (i.e., Novices) could more confidently be distinguished from higher ones (i.e., Experts and Intermediates). However, a separation of the higher skill categories was more difficult because the means of both categories in various criteria did not differ by more that both groups' standard deviation/variance.

Consequently, I was confident that my quantitative criteria could be used effectively to analyse archaeological assemblages and discriminate between bifaces that exhibited skilled (i.e., Expert or Intermediate) and unskilled (i.e., Novice) manufacture at GZnivG. Nevertheless, I understood that their power to discriminate between Expert and Intermediate in the archaeological record was reduced and that my quantitative methodology would need to be improved so that the differences between experienced knappers (i.e., Experts and Intermediates) that were revealed with qualitative assessment, could be also evinced quantitatively.

Qualitatively, experimental Intermediates made more frequent mistakes, such as failed flake terminations, and could not achieve similar levels of edge thinness and sharpness as their contemporary Expert counterparts. In fact, the only non-Expert experimental biface which truly resembled Expert handaxes in qualitative terms was CS_LIV2.13 which, significantly, was made under Expert instruction. Experts did not always perform consistently, however, and exhibited both quantitative and qualitative indicators of poorer

performance (e.g., biface KL_LIV.20). Reasons for why this may have occurred were identified as: poorer quality raw material, greater expediency employed in manufacture, or dips in performance by an Expert (Gray and Lindstedt, 2017). Another cause may have been that the only instruction they were given was to make an Acheulean-style handaxe; they were under no obligation to perform to their very highest standards and may only have been concerned with making a serviceable edge.

Ultimately, Experts and Intermediates exhibited a very similar range of scores which, counter to my expectations. made it difficult to separate them quantitatively. Occasional dips in performance by Experts (e.g., biface 'KL_LIV.20'), as well as unanticipated leaps in improvement among Intermediates (e.g., biface 'CS_LIV2.13') compounded this further. In addition, although Novices (especially Experimental knappers 6 and 7) performed consistently less well than Experts and Intermediates and could be distinguished more easily, examples such as 'LH_PSTKL.1', made by Experimental knapper 7, showed how significant the effect step-by-step instruction had on discriminating between knappers of varying experience. This has ramifications for correctly identifying the skill-level of knappers in the archaeological record by bifaces alone. Indeed, depending on circumstances of manufacture (i.e., with/without tuition), the same individual can produce two handaxes that are quantitatively and qualitatively very different.

Nevertheless, Experimental (Novice) knapper 4 showed that individuals technically considered as 'Novice' could overlap with some of the poorer Expert quantitative scores without step-by-step instruction. This was another noteworthy outcome which brought

into question the rationale behind the skill-level boundaries I used, and further questioned the effectiveness of some of my quantitative criteria.

To recap, Experimental knapper 4's 'Novice' bifaces were generally superior to other bifaces made by Novice knappers in my experimental investigation. Indeed, three of this knapper's bifaces which were made in their first year of practice were quantitatively misclassified as 'Expert' ('CS_LIV1.1', 'CS_LIV1.4', and 'CS_LIV1.6'). Although qualitative analysis revealed that these bifaces exhibited reduced levels of invasive flaking, less even thinning in profile view, a lack of post-thinning edge regularisation, and thicker and duller edges versus Expert bifaces in general, they were qualitatively more like Expert bifaces than other Novice (Experimental knappers 6 and 7) bifaces.

Experimental knapper 4's performance thus raises the question of innate ability and suggests my rationale for placing the Novice threshold at 1 year should potentially be revised. Based on Pargeter et al.'s (2019) finding that circa 200 hours of practice was needed to develop the bodily control required to reliably detach large thinning flakes, I estimated that one year would be the minimum amount of time to amass this amount of practice and graduate novicehood. Yet, it is clear via qualitative analysis that Experimental knapper 4 developed this ability within their first year of practice and possessed more than a Novice command of bifacial manufacture. Additionally, Novice Experimental knapper 7 showed significant leaps in improvement with practice and tuition over a period of 6 months which resulted in quantitative scores and qualitative assessment more like bifaces occasionally made by Intermediate participants.

Overall, the quantitative and qualitative overlaps between some of the bifaces made by Novice Experimental knappers 4 and 7 and some of the bifaces made by Experts and Intermediates, suggests that the threshold of novicehood should potentially be brought forward when conducting future experimental investigations. Although assigning arbitrary boundaries to skill-levels is somewhat problematic because development is quicker for some than others and is also occasionally influenced by other factors like tuition and/or varying amounts of practice, establishing skill-level thresholds was necessary in this investigation to reveal the variability between individuals with similar amounts of experience. Nonetheless, Experimental knapper 4 (via practice) and Experimental knapper 7 (via tuition) signalled that true novicehood was potentially relatively short in the Acheulean.

Direct instruction, in addition to imitation, was likely part of the Acheulean learning scenario (Lycett, 2019; Morgan et al., 2015; Shipton, 2019). Consequently, alongside necessary repeated practice, tuition's capacity to help expedite skill development may explain the rarity of true Novices at sites like GZnivG. We must remember, however, that learning in the Acheulean would likely have involved children and not the adults that participated in my experiment. Differences in strength and manual dexterity may have affected the speed with which knapping skill was developed among younger individuals in the Palaeolithic.

Certainly, though, my experimental investigation would have benefitted from a larger sample of bifaces from more participants. The restricted number of participants and bifaces is a potentially limiting factor in assessing the quantitative overlaps between skill-levels, particularly between Intermediates and Experts. Moreover, greater consideration

might have been given to the effect the method of biface manufacture had on my quantitative criteria. As Novice biface 'CS_LIV4' plainly demonstrated, increased refinement scores commensurate with Expert/Intermediate levels can be achieved via producing a biface on a large flake. Unsurprisingly, this bias extended to the classification of GZnivG bifaces, with bifaces like GZ2802(15) made on flakes being optimistically misclassified as Expert/Intermediate when qualitative assessment showed no indications of skilled work. However, although my mixed method has not produced a quantitatively complete way to distinguish between all skill-levels, it has provided a coherent approach which allows for ambiguities to potentially be resolved through correlation of priorly, and independently made, qualitative observations.

The main bias inherent in my experimental quantitative data when applying it to the classification of archaeological bifaces was the absence of resharpening employed by experimental knappers. Unfortunately, due to time and other constraints, this could not be mitigated. Among my experimental assemblage, more highly skilled individuals generally produced larger, more refined bifaces than lower skilled individuals. This appreciably affected accurate classification of GZnivG bifaces by skill-level category because, as McPherron (1999) noted, resharpening affected the refinement scores a biface could achieve below a certain length threshold. However, smaller bifaces, as my qualitative/quantitative classification comparison showed (see section 7.4.2.1), did not necessarily exhibit more indicators of poor performance. Predictably, this led to misclassifications of smaller GZnivG bifaces as Novice or Intermediate/Novice.

This result suggested that my quantitative criteria would be more accurate among a biface assemblage where resharpening had not occurred. However, for the comparison of skill-

level differences to be effective, it was necessary that the assemblage still demonstrated a coherent, shared idea of biface manufacture, like GZnivG. To test this idea further I examined a randomised sample of 57 bifaces from Boxgrove in section 8.3 below. These Boxgrove handaxes reveal evidence of tranchet sharpening but do not exhibit signs of re-sharpening and, despite differences in initial raw material size and shape and reduction sequences, yield very similar ultimate shapes (Garcia-Medrano et al., 2019; Leroyer, 2016; Pope, 2002; Roberts and Parfitt, 1999).

Before I present this analysis, however, I conclude this section with one of the most encouraging findings from my thesis. The Blind Test results, although based on a very small sample size, revealed that Experts were more consistently able to reproduce a handaxe to a specified design. As a result, when constraints on manufacture were introduced, it was easier to distinguish quantitatively between Experts and Intermediates than it was in the main study. Indeed, this result supports questionnaire participants' 5, 6, and 8 comments (see Catalogue 2, Q.17) that competence in a knapper can be boiled down to their capacity to replicate a biface. This is because replication of a standard, in relation to variations in raw material quality, primarily requires greater flexibility and predictive abilities regarding one's intentions. Individuals with greater experience possess a broader repository of sensory consequences of action and can thereby predict more accurately (even subconsciously) such intentions. They are in a greater state of 'action-readiness', allowing them to deal more effectively with varying material affordances and emerging intentions, so that their prior, conscious intentions (i.e., replicating an example biface) are achieved with high fidelity (Bruineberg et al., 2018; Bruineberg and Rietveld, 2014; Friston, 2011, 2010, 2009; Friston et al., 2011, 2009, 2010; Rietveld and Kiverstein, 2014).

Accordingly, imposing constraints on manufacture is a potentially fruitful avenue for more effectively quantitatively distinguishing between Experts and Intermediates, both experimentally and in the archaeological record. Social constraints on handaxe manufacture are visible in the Acheulean record at Boxgrove (García-Medrano et al., 2019; Hutchence and Scott, 2021; Leroyer, 2016; McNabb, 2017; Shipton and Clarkson, 2015; Shipton and White, 2020) and at GZnivG too. Therefore, morphological constraint is arguably a relevant aspect of biface manufacture in the Acheulean that was missing from my main experimental investigation. Thus, the variability of morphological, and other, intentions in my experimental data may have served to reduce the differences between Experts and Intermediates. Nevertheless, as discussed, the effects of biases like resharpening still need to be acknowledged and accounted for.

Later, in section 8.4, I will outline how the above insights could influence future methods for investigating skill-level differences in the Acheulean. First, however, I will assess whether my methodology is more accurate when applied to a non-resharpened assemblage from Boxgrove.

8.3 Further testing of my methodology – Boxgrove, UK

Here, I present a discriminant analysis based on Boxgrove handaxe data that has been normalised in line with the experimental LDA 1 previously used to analyse GZnivG bifaces. I follow this by outlining my observations of the post-LDA classification univariate statistics. The results of a qualitative analysis of the skill-level exhibited by the

bifaces are then correlated with the quantitative classifications. All Boxgrove data was collected via the British Museum's Franks House and can be found in Catalogue 6.

8.3.1 Boxgrove LDA

With the thresholds for skill-level outlined in section 7.4.1.2, I made a provisional assessment of the skill-level of the BXG handaxes based on their quantitative criteria scores:

Skill-level	Total
Novice	3
Intermediate/Novice	0
Intermediate	1
Expert/Intermediate	26
Expert	27
Total	57

Table 8. 1. Initial quantitative skill-level classification of Boxgrove sample.

To account for human bias, I performed a LDA to check the accuracy of my initial classifications.

- Jackknifed accuracy is 94.74%
    - 1 Novice biface was jackknifed to Intermediate.
    - The only Intermediate biface was jackknifed to Expert/Intermediate.

- 1 Expert/Intermediate was jackknifed to Intermediate.
- Expert classification was jackknifed as 100% accurate.

Loadings

	Axis 1 (96.61%)	**Axis 2 (3.36%)**	**Axis 3 (0.03%)**
Axis 1 (normalised)	81.343	4.0401	7.1081
Axis 2 (normalised)	-3.1378	9.6736	65.216
Axis 3 (normalised)	-6.4417	-47.77	-11.82

Table 8. 2. Loading scores by axis of normalised Boxgrove axes.

Confusion Matrix

	Novice	**Intermediate**	**Expert/Intermediate**	**Expert**	**Pre-LDA Total**
Novice	2	1	0	0	3
Intermediate	0	0	1	0	1
Expert/Intermediate	0	1	25	0	26

Expert	0	0	0	27	27
Post-LDA Total	2	2	26	27	57

Table 8. 3. Confusion matrix of Boxgrove sample bifaces.

Scatter plot

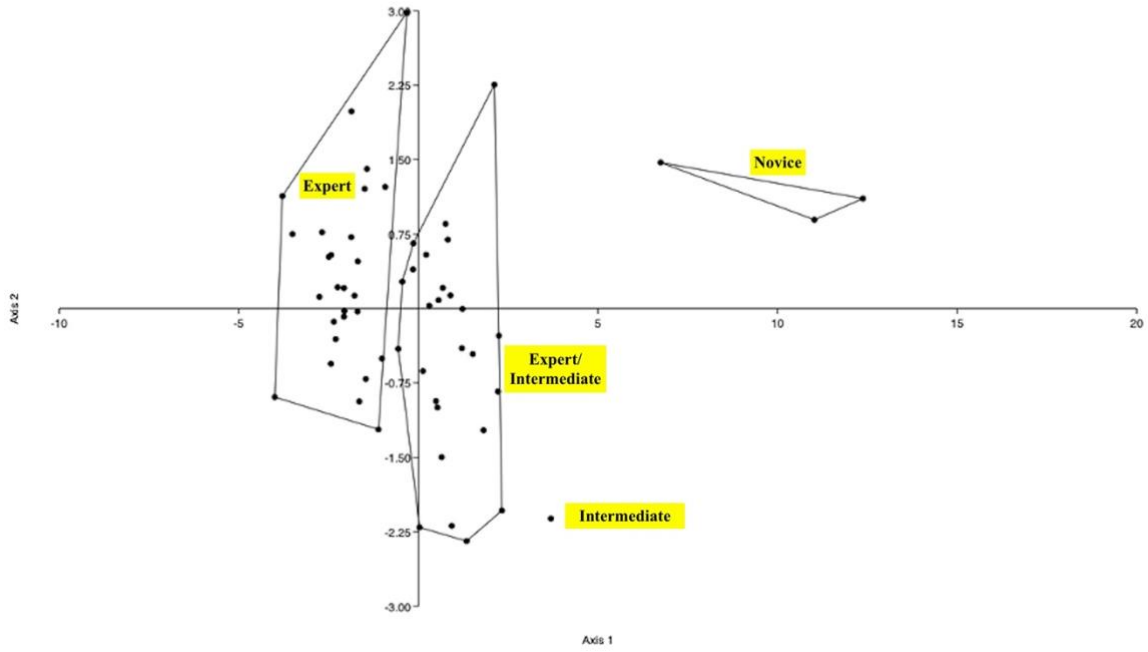


Figure 8. 1. LDA scatter plot of Boxgrove bifaces by initial classification.

It is clear from the initial classification scatter plot that Novices are separated markedly from other skill-levels. Table 8.3 demonstrates that initial classifications were very accurate, but LDA helped to further enhance the accuracy of these initial classifications (see Table 8.1) by resolving the human error inherent in them.

In passing it should be noted that this is a randomised assemblage of Boxgrove bifaces. Thus, my analysis of skill-level classification at the site does not provide a clearer

indication of the cross-section of skill-level categories involved at a particular locale. As such, I cannot comment on the proportion of skill-levels at such localities, as Leroyer (2016) has done elsewhere. My main aim, in relation to Boxgrove, is to assess whether my quantitative methodology is better at correctly classifying skill-level among a principally non-resharpened biface assemblage than it was for GZnivG's resharpened assemblage.

In fact, the Boxgrove LDA results clearly suggest this to be the case. For instance, the jackknifed BXG classification accuracy (94.74%) is far greater than the 69.77% accuracy at GZnivG (see Table 7.41). In addition, the significant quantitative over-classification of lower-skilled bifaces at GZnivG does not seem to have occurred in this Boxgrove sample. Indeed, whereas c.30% (n = 39) of GZnivG bifaces were jackknifed as Novice or Intermediate/Novice, only c. 4% (n = 2) of BXG bifaces were classified to these categories.

The smaller number of Novice and Intermediate/Novice classifications at BXG could, however, have been a result of the randomised nature of the Boxgrove sample. Thus, to properly determine that the lack of resharpening at BXG was the predominant reason for greater classification accuracy, it was still necessary to double check that elongation bias, or any other bias (e.g., raw material quality), was not affecting the above LDA classifications. Subsequently, in the next section I present my observations of the post-LDA univariate statistics (see Catalogue 6, Sheet 6.3), which is followed by a qualitative analysis of BXG bifaces.

8.3.2 Post-LDA univariate statistics

It is important to note that when assessing the significance of certain univariate statistics, sample size for each skill-level category should be considered carefully. An illustration of this is that both Novice and Intermediate classified bifaces only amount to a very small proportion of the overall assemblage. This needs to be remembered when attempting to interpret the presence of biases in the data.

For example, at first glance, the post-LDA univariate statistics for Length and Breadth (see Catalogue 6, Sheet 6.3, Tables 1 and 2) suggest that biface size shows a positive bias for Experts as at GZnivG. The minimum and maximum lengths, as well as the means, suggested that bifaces assessed as being crafted by Experts were larger than those assigned to lower skill-levels. In fact, only two Expert handaxes were under the mean length for the entire assemblage (125.38mm). However, the sample sizes of lower-skill level categories (Intermediate and Novice) in the Boxgrove case were too small to be fairly compared with the majority of bifaces that were classified as Expert and Expert/Intermediate. Moreover, the Expert/Intermediate category shows that smaller bifaces were categorised as higher skill-levels too. Indeed, the two smallest lengths and breadths are in Expert/Intermediate and Intermediate categories and not attributed to Novice. Therefore, it was difficult to attribute any size bias to the classification of BXG bifaces.

For skill related criteria (i.e., Ratio of Refinement, Thickness/Length ratio, Planform Asymmetry, and Edge Straightness), as well, the sample sizes for bifaces classified as

Novice and Intermediate were generally too small to make more concrete observations other than higher skill-levels produced better scores. Of these criteria, Thickness/Length ratio and Edge Straightness data exhibited the least intersection between the skill-level categories. This mirrors the experimental data used to classify via LDA. Therefore, although post-LDA statistics suggested that length/elongation bias may not have been as strong among mostly non-resharpened bifaces, qualitative analysis was needed to make sure this was the case given the small sample sizes involved.

8.3.3 Qualitative analysis

As explained in Chapter 7, my qualitative classifications for GZnivG were considered 'correct' if they were within 1 place, or skill-level denomination, of the 5 quantitatively derived skill-levels. Fuller detail on qualitative observations for all bifaces listed in Table 8.4 below is provided in Catalogue 6 (Sheet 6.4).

Generally, qualitative analysis also worked well on the BXG sample and supports the quantitative skill-level classifications given to each biface. Only 3 out of 57 qualitative assessments fundamentally disagreed with quantitative classifications Some of the key examples that are included below show where qualitative analysis either supports or contradicts quantitative classification. Moreover, I have included examples that demonstrate the key contextual information that qualitative assessment provides when quantitative classification is borderline (e.g., Expert/Intermediate assignations).

Biface ID	Given group	Jackknifed	Qualitative correlation (Correct/Incorrect)
84	Expert	Expert	✓
86	Expert	Expert	✓
89	Expert	Expert	✓
90	Expert/Intermediate	Expert/Intermediate	✓
92	Expert/Intermediate	Expert/Intermediate	✓
95	Expert/Intermediate	Expert/Intermediate	✓
97	Expert/Intermediate	Expert/Intermediate	✓
98	Expert	Expert	✓
116	Expert	Expert	✓
119	Expert/Intermediate	Expert/Intermediate	✓
121	Expert	Expert	✓
122	Expert	Expert	✓
124	Expert	Expert	✓
127	Expert	Expert	✓
128	Novice	Novice	✓
131	Expert	Expert	✓
132	Expert/Intermediate	Expert/Intermediate	✓
153	Expert/Intermediate	Expert/Intermediate	✓
155	Expert	Expert	✓
156	Expert/Intermediate	Expert/Intermediate	✓
159	Expert/Intermediate	Expert/Intermediate	✓
160	Expert/Intermediate	Expert/Intermediate	✓
163	Expert	Expert	✓
164	Expert/Intermediate	Expert/Intermediate	✓
167	Expert/Intermediate	Expert/Intermediate	✓
168	Expert	Expert	✓
204	Expert/Intermediate	Expert/Intermediate	✓

207	Expert	Expert	✓
209	Expert	Expert	✓
211	Expert	Expert	✓
212	Expert/Intermediate	Intermediate	✓
214	Expert	Expert	✓
216	Novice	Intermediate	×
218	Expert/Intermediate	Expert/Intermediate	✓
220	Expert/Intermediate	Expert/Intermediate	✓
222	Expert/Intermediate	Expert/Intermediate	✓
224	Expert/Intermediate	Expert/Intermediate	✓
226	Expert/Intermediate	Expert/Intermediate	✓
229	Expert	Expert	✓
231	Expert	Expert	✓
232	Novice	Novice	✓
264	Expert	Expert	✓
267	Expert	Expert	✓
269	Expert	Expert	✓
270	Expert	Expert	✓
272	Expert/Intermediate	Expert/Intermediate	✓
274	Expert	Expert	✓
277	Expert	Expert	✓
279	Expert	Expert	✓
281	Expert	Expert	✓
283	Intermediate	Expert/Intermediate	×
284	Expert/Intermediate	Expert/Intermediate	✓
286	Expert/Intermediate	Expert/Intermediate	✓
288	Expert/Intermediate	Expert/Intermediate	✓
291	Expert/Intermediate	Expert/Intermediate	✓
292	Expert/Intermediate	Expert/Intermediate	✓

294	Expert/Intermediate	Expert/Intermediate	✓

Table 8. 4. Qualitative corroboration of quantitative skill-level assessments of Boxgrove bifaces.

Boxgrove biface examples that show good fit with quantitative classification:

- Biface 98
    - Relatively small ovate with continuous cutting edge (and tranchet blow). Displays what looks like considerable cortex cover but was in fact chalk inclusions which would have made production of this thin, well-made tool more difficult.
- Biface 128
    - Resembles the work of a lower skilled individual producing a roughout on a small nodule. The biface is covered in cortex and incipient cones which show the knapper was hitting it too hard and at the wrong angle. There are inclusions but as other bifaces (e.g., 98) show, inclusions have not prevented other more skilled knappers from effectively creating a handaxe).
- Biface 212
    - Small, pointed form with irregular shape. It exhibits a poorly worked edge which is sinuous and zig-zagged with some cortex visible. The biface's symmetry is reduced also. I suggest this biface resembles the lower end of Intermediate-level knapping. This nicely corroborates its quantitative classification given that the 'Intermediate' range relates to the lower end of range of intermediate scores in my experimental assemblage (as outlined in Chapter 7).

- Biface 232
    - Completely but irregularly knapped nodule. There is no clear plan, and the knapper was unable to manufacture a continuous cutting edge. The biface is a small and thick sub-cordate which looks more like a core for flake production. Overall, it is poorly made.
- Biface 291
    - Small, pointed form made on a thin nodule (cortex on both faces). I suggest this is expediently made and not poorly made, given the thinness of the nodule. This is supported by well-maintained edge, despite not all of the cortex being removed. My interpretation suggests an experienced knapper utilised an already thin nodule to make a biface quickly.

<u>Boxgrove biface examples where qualitative assessment provides contextual support to borderline quantitative classification:</u>

- Biface 90
    - Smaller, pointed ovate. Thicker on one face but exhibits skilful knapping including good edge maintenance and the successful removal of a tranchet flake. Qualitative analysis suggests its boundary quantitative classification is fair.
- Biface 95
    - Small ovate that is more irregular shape. There is some cortex remaining. Moreover, a large removal at the base of the biface leaves an irregular edge. Additionally, steeper flake removal scars are indicative of poorer

thinning, but may also relate to the size of the tool. Qualitative analysis suggests the knapper was more Intermediate than Expert.

- Biface 132
    - Small, more irregularly shaped elongated ovate. This biface's continuous cutting edge is sometimes interrupted by cortex towards its butt. The knapper successfully removed a tranchet flake and produced a thin biface. Planform symmetry was reduced, however, and evidence of hinge fractures are visible. Qualitative assessment would therefore suggest the maker of this biface is more likely Intermediate.
- Biface 159
    - Very small circular ovate. The knapper has produced a continuous cutting edge but no tranchet flake. This is likely to do with the size of the biface. Otherwise, it is very finely made. The knapper was able to make a very thin biface that is also small, whilst maintaining a straight edge. This is a difficult skill to master and suggests that even when resharpening is not an issue, elongation bias among smaller bifaces can influence classification. Although this biface is classified as Expert/Intermediate, it is clear from qualitative assessment that this was made by a highly skilled practitioner.
- Biface 294
    - Small, pointed form made on a very thin nodule which is demonstrated by there being cortex visible on both faces. As suggested for biface 291, this biface looks to be expediently made. Nevertheless, qualitatively this biface is well made, indicated by good edge maintenance.

Boxgrove biface examples where qualitative assessment contradicts quantitative classification:

- Biface 216
    - Small, irregular pointed form which does not exhibit a straight cutting edge. One edge is curved due to thick, steep removals that have not effectively thinned the biface. In addition, cortex located on the central ridge of one face also suggests the knapper was unable to thin effectively and/or remove all the cortex. Plano-convexity in profile view may also indicate this biface was made on a flake. As I demonstrated above, increased refinement scores can be achieved when a biface is made on a flake. This may explain how this biface is qualitatively assessed to be made by a Novice but yields a quantitative Intermediate classification.
- Biface 283
    - Very small ovate pointed biface with an irregular shape. This biface's cutting edge is irregular and, as such, it is difficult to determine proximal from distal ends. Step/hinge fracture scars are visible on both faces. My initial classification based on thresholds put this biface as Intermediate, however, it was jackknifed to Expert/Intermediate. However, edge maintenance and thinning are poor, and I would overturn the jackknifed classification and qualitatively classify this biface as Novice, especially considering this biface has the third lowest edge straightness score and bottom 20% of Th/L scores (the two strongest loadings in the LDA used to classify this, and all, bifaces).

Overall, I could not find bifaces in my sample overly affected by length/elongation bias in their quantitative classification. Smaller handaxes that were qualitatively observed to be well, or very well, made (e.g., Bifaces 159 and 224), were most commonly classified as Expert/Intermediate, however. This suggested that length was still having a slight detrimental influence on these bifaces not reaching the quantitative threshold for Expert.

For bifaces 159 and 224, Th/L ratio and Edge Straightness were the main factors affecting their quantitative classification. Both bifaces' Ratio of Refinement and Planform Asymmetry scores were below the Expert mean, whereas their Th/L ratio and Edge Straightness scores were between the Expert and Expert/Intermediate means. Both criteria use biface length in their calculation (see section 6.4.1, Chapter 6). They were also the two strongest loadings used to classify archaeological bifaces via LDA. This therefore explains why these highly symmetrical, well thinned, straight edged, and very skilled bifaces were 'marked down' quantitatively.

Consequently, I have demonstrated that the biasing effect that length has on classification of non-resharpened bifaces is not as severe as I originally thought based on the outcomes of the GZnivG assemblage outlined in chapter 7. The only notable source of quantitative misclassification in the above sample was when bifaces (e.g., Biface 216) occasionally achieved greater refinement scores through being made on a flake. This also occurred in the GZnivG assemblage (e.g., Bifaces GZ2802(15) and GZ2812(22)). Nevertheless, overall, my quantitative criteria have been proven to be much better equipped to accurately classify bifaces that exhibit little to no re-sharpening by the skill-level of their maker than resharpened ones.

In the next section, I reflect upon the outcomes of my thesis, identifying areas which have the potential for greater resolution and future directions of research.

8.4 Main outcomes of my research and their relation to areas of future study

My analysis of Novice Experimental knapper 7's bifaces showed that the quantitative criteria used can accurately predict when tuition has taken place. This supported the discussion in Chapter 4 that skill in these sessions is being borrowed and is not yet embodied. It also confirmed that skill-level cannot be transferred from one individual to another (Ingold, 2013; Keller and Keller, 1996). Experimental knapper 7's performance improved dramatically with tuition but unaided then fell back to near the levels reached pre-tuition. This appears to support Gray and Lindstedt's (2017) notion of PDLs, as well as the idea that the only way to improve is through practice and engagement with the material to develop greater flexibility and affinity with the material being worked.

Applying this finding to the archaeological record is of course much more challenging. Demonstrating that a single knapper was responsible for multiple bifaces is extremely problematic for most assemblages. In fact, even if perfect conditions were discovered, like the famed Upper Palaeolithic workshop sites of Pincevent (Ploux, 1989) and Etiolles (Pigeot, 1990), it would be very difficult to propose that two bifaces were made by the same individual if one was made with tuition and the other was made without it. Indeed, my comparison of two of Experimental knapper 7's bifaces (LH_KL.2 and LH_PSTKL.1) in Chapter 7 (see section 7.2.2) revealed very few similarities between them that would, without knowing they were made by the same individual, lead the

assessor to identify this correctly. Thus, reliably demonstrating that direct step-by-step instruction occurred in the Acheulean would require investigators to be able to accurately establish that two bifaces were made by the same individual and that one of them exhibits a marked improvement in performance. This is highly unlikely and, as a result, dependably determining the effect tuition has on inexperienced knappers is improbable in the archaeological record. Nevertheless, further experimental work on finding methods for identifying individuals when performing both with and without tuition would be a very useful area of future study that could, if successful, lead to conclusively demonstrating that direct, assisted learning occurred in the Acheulean.

Separately, the results of my Blind Test suggested that imposing restrictions on manufacture related to biface morphology may provide a promising avenue for elucidating differences between Expert and Intermediate knappers. Initial results presented here show unequivocally how Expert knappers more reliably reproduce a given artefact than a less experienced Intermediate. This is useful evidence when discussing the location and constraining of intentions, as discussed in Chapter 4 (Bruineberg and Rietveld, 2014; Hutchence and Scott, 2021; Malafouris, 2013; Pelegrin, 2005; Rietveld and Kiverstein, 2014). Further research using a much larger sample of bifaces will of course be required before the results can be more confidently accepted.

In the case of the knappers (Experimental knapper 1 and Experimental knapper 4) tasked with replicating the biface pictured in Figure 6.1, both knapped three handaxes. The Expert (Experimental knapper 1) achieved consistently similar quantitative scores to the example Master biface specimen. This points towards greater control by the Expert over the Intermediate regarding their incipient intentions in relation to the emerging

affordances of the material. Indeed, qualitative observation corroborated quantitative results. Therefore, my results imply that the Expert could better regulate the emerging intention/affordance scenario in line with priorly situated intentions that relate to replicating a particular example's morphology.

This discovery also possesses a broader significance given that similar constraints on manufacture arguably existed at Gouzeaucourt, Boxgrove, and many other Acheulean sites besides. Indeed, diacritical analysis presented in Chapter 7 shows that there were two predominant flake removal patterns employed at GZnivG. However, as I expected, bifaces categorised as being made by lower skill-levels behaved less consistently and did not employ the most common removal configurations of higher skill-levels. This potentially implies that the less proficient knappers at GZnivG, as per my experimental findings, found it harder to constrain their emerging intentions in line with the conventions constraining their performance.

Therefore, although I was unable to systematically reveal clear quantitative differences between Expert and Intermediate knappers in my main experimental investigation, my Blind Test results showed that there is the potential to achieve this in future experimental work. Key to this will be to incorporate a greater number of experimental participants, so that the resulting outcomes are more statistically robust. Consequently, if both tests (main experimental and blind) are repeated with larger sample sizes for each skill-level category, the abilities of variably experienced modern knappers can be compared when constraints are imposed or not. This will ultimately demonstrate whether imposing constraints on manufacture genuinely has a considerable impact on more effectively quantitatively discriminating between Expert and Intermediate bifaces, as my Blind Test

suggested. Moreover, it would provide an opportunity to properly test the utility of additional quantitative criteria (e.g., Flake Scar number) for discriminating between skill-levels.

However, the results of this future analysis may be particularly effective in analysing biface assemblages where resharpening appears absent. As we saw at GZnivG, despite arguable constraints on the morphology of bifaces produced at the site, elongation bias would have likely precluded accurate classification even with a larger Blind Test sample. Hence, I suggest that a useful future investigation would be to find a better way to both quantitatively and qualitatively assess skill-level differences among resharpened assemblages. I do not have the scope here to detail in full how this should be approached, however I envisage that it must try to replicate the patterns of resharpening visible among the specific bifaces being analysed. Using GZnivG as an example, this would necessitate knappers of varying abilities attempting to replicate the flaking/resharpening techniques highlighted by Soriano (2000). This will be extremely challenging and require a great deal of time, participants, resources, and planning. But it would be worthwhile to see if differences between the skill-levels could be identified when varying levels of resharpening has occurred.

Overall, in line with the main broader aim of my thesis, future experimental studies should be focused on improving the accuracy with which the skill-level of the maker can be quantitatively identified. I have demonstrated that my quantitative criteria are successful in distinguishing between contemporary knappers of higher and lower skill-levels (particularly for ovate handaxes), but future research should strive to reduce the quantitative overlap between experimental Experts and Intermediates that has been

presented in this thesis. Doing so will increase the overall confidence in applying experimental data to analyse skill-level in the archaeological record and thus provide new data to aid interpretation of site function in the Acheulean – as I have done at GZnivG.

8.5 Assessing site function

A crucial outcome of my thesis is that, by assessing skill-level differences among an archaeological biface assemblage, I have generated data regarding the proportions of skill-categories active at that site. This information can be used as a proxy for the experience and relative ages of the knappers that contributed to the assemblage.

My mixed methodology that combined quantitative and qualitative analysis demonstrated that Novices were least represented skill-level category at GZnivG. After assessing the impact of resharpening on quantitative classifications, only 16 (out of 129) (see Catalogue 3, Sheet 3.1) were considered to have exhibited Novice-level performance. It must be reiterated that these do not represent individuals per se, but categories of individuals grouped by their knapping abilities.

The scarcity of Novice presence at GZnivG validates part of Hypothesis 11, which reasoned that because Novices and Experts constitute the extremes of the skill spectrum, by default they will be rarer than Intermediates. But counter to Hypothesis 11, the results outlined in Chapter 7 show that there were proportionally many more instances of very skilled performers at GZnivG. The significance of this result resides in its support for

previous interpretations of GZnivG (Lamotte, 1992; McPherron, 1994; Scott and Ashton, 2010; Soriano, 2000; Tuffreau et al., 2008; Tuffreau and Bouchet, 1985).

As detailed in Chapter 5, GZnivG bifaces were produced off site (McPherron, 1994; Scott and Ashton, 2010; Tuffreau et al., 2008). Reduction and re-sharpening analysis, as well an absence of evidence of refitting between flake debris that was found in association with the bifaces, attests to this (McPherron, 1994; Scott and Ashton, 2010; Tuffreau et al., 2008). In addition, bifaces occasionally showed secondary signs of modification with flake removals more consistent with later prepared-core technologies (Lamotte, 1992; Soriano, 2000; Tuffreau and Bouchet, 1985). This further demonstrates that post-depositional factors were not the principal cause of movement of stone tools into and out of GZnivG. As Tuffreau et al. (2008) explain, bifaces were being curated and regularly moved across the landscape by hominins in MIS 8, the date attributed to GZnivG. If correct, it may imply that assemblages like GZnivG's were the result of hominins intentionally conserving raw material in an increasingly severe climatic environment (Hérisson et al., 2016; Malinsky-Buller, 2016; Scott and Ashton, 2010).

In consequence, it is likely that the GZnivG assemblage was formed over multiple visits by groups with a shared tradition of biface manufacture. These groups left bifaces behind which subsequently acted as a pre-prepared synthetic outcrop, or cache, for later visits by their own or other groups. This interpretation of movement in and out of the site is supported by the high numbers of handaxes which were capable of being resharpened and used further.

Therefore, GZnivG must have been a significant place in the landscape for hominins. Indeed, there are other levels (i.e., H and R) exhibiting handaxe assemblages too. Moreover, secondary modifications of bifaces more consistent with later prepared-core technologies at GZnivG suggests this location was visited repeatedly over time by groups with different technological repertoires (Lamotte, 1992; Soriano, 2000; Tuffreau and Bouchet, 1985). It may even suggest that the positions of such stone raw material cache locations were transmitted between groups over large distances.

Therefore, it follows that with the changes in relationship to material during the transition from MIS 9 to 8 and 7 (more resharpening and conserving of raw material) (Hérisson et al., 2016; Malinsky-Buller, 2016; Scott and Ashton, 2010), it would be less likely to find evidence of poor craftspersonship in transitional sites like GZnivG where handaxes were moved over significant distances. In Hypothesis 11, I reasoned that Novices would be less likely to carry their poorly made bifaces with them due to them likely corresponding to younger individuals with reduced carrying capacity. Thereby, bifaces categorised as being made by Novices would be fewer in number at Gouzeaucourt. Alternatively, the pattern from GZnivG could indicate that less skilled individuals were prohibited from working on valuable raw material that was intentionally conserved due to the scarcity of immediately available high quality raw material. Instead, the Novices that belonged to these groups may be more prevalent among assemblages near quarry locations. Moreover, as I postulated earlier in this chapter, novicehood could potentially be a relatively short period in the lives of Acheulean knappers. If this was the case, then the chances of finding Novice behaviour will always be reduced.

Consequently, conducting the same investigation presented here on the biface assemblages from other levels at Gouzeaucourt would be informative regarding whether they yield similar paucity of Novices too. Bifaces found in Levels H and R differ to those in G (GZnivG) in that they were manufactured *in situ* (Scott & Ashton, 2010; McPherron 1994). Thus, Novices may potentially be more numerous among these assemblages than at GZnivG.

8.6 Conclusion

Discovering new ways of prising data from the inert stones our hominin ancestors left behind is a challenging but worthwhile pursuit. This thesis has been my attempt to contribute to this task. My contribution has demonstrated that determinations of skill-level in the archaeological record may be made by analysing experimental biface data. Importantly, my methodology can be widely applied; each of my quantitative and qualitative criteria can be calculated using equipment easily accessible to all researchers. I have shown that my quantitative criteria can accurately distinguish higher skill-levels from lower ones, especially among assemblages with low evidence of re-sharpening. Moreover, qualitative analysis is very effective at identifying biases (i.e., resharpening/elongation bias) one may suspect to exist in quantitative datasets. The combination of the two has provided me with greater confidence in my skill-level classifications.

However, as detailed above, further experimental work is now needed to isolate the best criteria for quantitatively separating Experts and Intermediates and reflecting the

qualitative differences between them that could be seen in my experimental investigation. Increasing sample size in future experiments and testing the effects of applying constraints to manufacture are key next steps in this pursuit.

My interpretation of skill-level at GZnivG suggests Novices were infrequent contributors to the activities at the site. This was likely due to the transitional nature of the site and the likelihood that Novices contributed comparatively fewer knapping products than more skilled individuals in the same group, as they travelled through the landscape. Overall, I consider my placement of GZnivG bifaces at different stages along a skill-spectrum to reasonably reflect the skill-level category of their makers.

Bibliography

Alperson-Afil, N., Richter, D., Goren-Inbar, N., 2018. Correction: Evaluating the intensity of fire at the Acheulian site of Gesher Benot Ya'aqov—Spatial and thermoluminescence analyses (PLoS ONE (2017) 12 (11) (e0188091) DOI: 10.1371/journal.pone.0188091). PLoS ONE 13. https://doi.org/10.1371/journal.pone.0190804

Alperson-Afil, N., Richter, D., Goren-Inbar, N., 2017. Evaluating the intensity of fire at the Acheulian site of Gesher Benot Ya'aqov - Spatial and thermoluminescence analyses. PLoS ONE 12. https://doi.org/10.1371/journal.pone.0188091

Alperson-Afil, N., Sharon, G., Kislev, M., Melamed, Y., Zohar, I., Ashkenazi, S., Rabinovich, R., Biton, R., Werker, E., Hartman, G., Feibel, C., Goren-Inbar, N., 2009. Spatial organization of hominin activities at Gesher Benot Ya'aqov, Israel. Science (New York, N.Y.) 326, 1677. https://doi.org/10.1126/science.1180695

Ameloot-Van Der Heijden, N., 1993. L'ensemble lithique du gisement de Longavesnes (Somme) : illustration d'un problème de reconnaissance du débitage levallois dans une industrie à bifaces de la phase ancienne du Paléolithique moyen. Bulletin de la Société préhistorique française 90, 257–264. https://doi.org/10.3406/bspf.1993.9591

Ashton, N., McNabb, J., 1994. Bifaces in Perspective, in: Ashton, N., David, A. (Eds.), Stories in Stone, Lithic Studies Society Occasional Papers 4. Lithics Studies Society, pp. 182–191.

Ashton, N., White, M.J., 2003. Bifaces and Raw materials: flexible flaking in the British Lower Palaeolithic., in: Dibble, H., Sorresi, M. (Eds.), Multiple Approaches to the Study of Bifacial Technologies. University of Pennsylvania Museum Press., pp. 109–124.

Austin, H.A., 1976. A computational theory of physical skill. Massachusetts Institute of Technology, Massachusetts.

Bae, C.J., 2013. Archaic Homo sapiens. Nature Education Knowledge 4, 4.

Bamforth, D., Finlay, N., 2008. Introduction: Archaeological Approaches to Lithic production skill and craft learning. Journal of Archaeological Method and Theory 15, 1–27.

Barham, L., Everett, D., 2021. Semiotics and the Origin of Language in the Lower Palaeolithic. Journal of Archaeological Method and Theory 28, 535–579. https://doi.org/10.1007/s10816-020-09480-9

Bates, M.R., Wenban-Smith, F.F., Bello, S.M., Bridgland, D.R., Buck, L.T., Collins, M.J., Keen, D.H., Leary, J., Parfitt, S.A., Penkman, K., Rhodes, E., Ryssaert, C., Whittaker, J.E., 2014. Late persistence of the Acheulian in southern Britain in an MIS 8 interstadial: evidence from Harnham, Wiltshire. Quaternary Science Reviews 101, 159–176. https://doi.org/10.1016/j.quascirev.2014.07.002

Berna, F., Goldberg, P., Horwitz, L.K., Brink, J., Holt, S., Bamford, M., Chazan, M., 2012. Microstratigraphic evidence of in situ fire in the Acheulean strata of Wonderwerk Cave, Northern Cape province, South Africa. Proc Natl Acad Sci USA 109, E1215. https://doi.org/10.1073/pnas.1117620109

Beyene, Y., Katoh, S., Woldegabriel, G., Hart, Uto, K., Sudo, M., Kondo, M., Hyodo, M., Renne, P.R., Suwa, G., Asfaw, B., 2013. The characteristics and chronology of the earliest Acheulean at Konso, Ethiopia. Proceedings of the National Academy of Sciences 110, 1584. https://doi.org/10.1073/pnas.1221285110

Biryukova, E.V., Bril, B., Frolov, A.A., Koulikov, M.A., 2015. Movement Kinematics as an Index of the Level of Motor Skill: The Case of Indian Craftsmen Stone Knapping. Motor Control 19, 34–59. https://doi.org/10.1123/mc.2013-0042

Bocquet-Appel, J.-P., Tuffreau, A., 2009. Technological Responses of Neanderthals to Macroclimatic Variations (240,000-40,000 BP). Human Biology 81, 287–307. https://doi.org/10.3378/027.081.0310

Boëda, E., 1995. Levallois: A volumetric construction, methods, a technique., in: Dibble, H.L., Bar-Yosef, O. (Eds.), The Definition and Interpretation of Levallois Technology, Monographs in World Archaeology N°23. Prehistory Press, Wisconsin, pp. 41–68.

Bordaz, J., 1971. Tools of the Old and New Stone Age, 1st ed. David & Charles, Newton Abbot.

Bordes, F., 1971. Physical evolution and technological evolution in man: A parallelism. World archaeology 3, 1–5. https://doi.org/10.1080/00438243.1971.9979488

Bordes, F., 1968. The Old Stone Age, World university library. Weidenfeld & Nicolson, London.

Botha, R., 2015. Teaching and Learning Subsistence Skills: Did Premodern Hominins Use Language to Do It? Cambridge Archaeological Journal 25, 901–908.

Bouchrika, I., 2021. How to Write Research Methodology: Overview, Tips, and Techniques. Guide2Research.com. URL https://www.guide2research.com/research/how-to-write-research-methodology (accessed 7.5.21).

Brannen, J., Moss, G., 2012. Critical Issues in Designing Mixed Methods Policy Research. American Behavioral Scientist 56, 789–801. https://doi.org/10.1177/0002764211433796

Bridgland, D.R., White, M.J., 2015. Chronological variations in handaxes: patterns detected from fluvial archives in north-west Europe. Journal of Quaternary Science 30, 623–638.

Bridgland, D.R., White, M.J., 2014. Fluvial archives as a framework for the Lower and Middle Palaeolithic: patterns of British artefact distribution and potential chronological implications. Boreas 43, 543–555.

Bril, B., Rein, R., Nonaka, T., Wenban-Smith, F., Dietrich, G., 2010. The role of expertise in tool use: Skill differences in functional action adaptations to task constraints. Journal of Experimental Psychology: Human Perception and Performance 36, 825–839.

Bruineberg, J., Kiverstein, J., Rietveld, E., 2018. The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. Synthese 195, 2417–2444. https://doi.org/10.1007/s11229-016-1239-1

Bruineberg, J., Rietveld, E., 2014. Self-organization, free energy minimization, and optimal grip on a field of affordances. Frontiers in human neuroscience 8, 599. https://doi.org/10.3389/fnhum.2014.00599

Bruner, E., Iriki, A., 2016. Extending mind, visuospatial integration, and the evolution of the parietal lobes in the human genus. Quaternary international 405, 98–110. https://doi.org/10.1016/j.quaint.2015.05.019

Callahan, E., 1979. THE BASICS OF BIFACE KNAPPING IN THE EASTERN FLUTED POINT TRADITION: A MANUAL FOR FLINTKNAPPERS AND LITHIC ANALYSTS. Archaeology of Eastern North America 7, 1–180.

Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science.(Editorial)(Author abstract). Behavioral and Brain Sciences 36, 181. https://doi.org/10.1017/S0140525X12000477

Clark, A., 2007. Re-inventing ourselves: the plasticity of embodiment, sensing, and mind. The Journal of medicine and philosophy 32, 263. https://doi.org/10.1080/03605310701397024

Clark, G., 1969. World prehistory : a new outline, 2nd ed. Cambridge University Press, Cambridge.

Coolidge, F.L., Wynn, T.G., 2009. The rise of homo sapiens : the evolution of modern thinking. Wiley-Blackwell, Chichester.

Coppe, J., Rots, V., 2017. Focus on the target. The importance of a transparent fracture terminology for understanding projectile points and projecting modes. Journal of Archaeological Science: Reports 12, 109–123. https://doi.org/10.1016/j.jasrep.2017.01.010

Corbey, R., 2020. Baldwin effects in early stone tools. Evolutionary Anthropology: Issues, News, and Reviews 29, 237–244. https://doi.org/10.1002/evan.21864

Corbey, R., Jagich, A., Vaesen, K., Collard, M., 2016. The acheulean handaxe: More like a bird's song than a beatles' tune? Evolutionary Anthropology: Issues, News, and Reviews 25, 6–19. https://doi.org/10.1002/evan.21467

Coward, F., Gamble, C., 2010. Metaphor and Materiality in Earliest Prehistory., in: Malafouris, L., Renfrew, C. (Eds.), The Cognitive Life of Things: Recasting the Boundaries of the Mind. McDonald Institute for Archaeological Research, Cambridge, pp. 39–46.

Crompton, R.H., Gowlett, J.A.J., 1993. Allometry and multidimensional form in Acheulean bifaces from Kilombe, Kenya. Journal of human evolution 25, 175–199. https://doi.org/10.1006/jhev.1993.1043

Darmark, K., 2010. Measuring skill in the production of bifacial pressure flaked points: a multivariate approach using the flip-test. Journal of Archaeological Science 37, 2308–2315. https://doi.org/10.1016/j.jas.2010.04.004

Davidson, I., Noble, W., 1993. Tools and language in human evolution, in: Gibson, K., Ingold, T. (Eds.), Tools, Language and Cognition in Human Evolution. Cambridge University Press, Cambridge, pp. 363–388.

Davidson, I., Noble, W., 1989. The Archaeology of Perception: Traces of Depiction and Language [and Comments and Reply]. Current Anthropology 30, 125–155. https://doi.org/10.1086/203723

de La Torre, I., 2016. The origins of the Acheulean: past and present perspectives on a major transition in human evolution. Philosophical Transactions of the Royal Society B 371. https://doi.org/10.1098/rstb.2015.0245

de Mortillet, G., 1873. Classification des diverses périodes de l'âge de la pierre., in: Extrait Du Compte Rendu Du Congrès International d'Anthropologie et d'Archéologie Préhistoriques, 6me Session, Bruxelles, 1872. M. Weissenbruch, Bruxelles, Belgium, pp. 432–444.

Degenaar, J., O'Regan, J., 2017. Sensorimotor Theory and Enactivism. Topoi 36, 393–407. https://doi.org/10.1007/s11245-015-9338-z

Dibble, H., Holdaway, S., Lin, S., Braun, D., Douglass, M., Iovita, R., McPherron, S., Olszewski, D., Sandgathe, D., 2017. Major Fallacies Surrounding Stone Artifacts and Assemblages. J Archaeol Method Theory 24, 813–851. https://doi.org/10.1007/s10816-016-9297-8

Dibble, H.L., 1984. Interpreting typological variation of Middle Paleolithic scrapers: Function, style, or sequence of reduction? Journal of Field Archaeology 11, 431–436.

Dibble, H.L., Pelcin, A., 1995. The Effect of Hammer Mass and Velocity on Flake Mass. Journal of archaeological science 22, 429–439. https://doi.org/10.1006/jasc.1995.0042

Dominguez- Rodrigo, M., Serrallonga, J., Juan-Tresserras, J., Alcala, L., Luque, L., 2001. Woodworking activities by early humans: a plant residue analysis on Acheulian stone tools from Peninj (Tanzania). Journal of Human Evolution 40, 289–299. https://doi.org/10.1006/jhev.2000.0466

Edwards, S.W., 2001. A modern knapper's assessment of the technical skills of the Late Acheulean biface workers at Kalambo Falls, in: Clark, J.D. (Ed.), Kalambo Falls Prehistoric Site, The Earlier Cultures: Middle and Earlier Stone Age. Cambridge University Press, Cambridge, pp. 605–611.

Emery, K., 2010. A re-examination of variability in handaxe form in the British Palaeolithic. UCL PhD Dissertation, London.

Eren, M.I., Bradley, B.A., Sampson, G.C., 2011. Middle paleolithic skill level and the individual knapper: An experiment. American Antiquity 76, 229–251. https://doi.org/10.7183/0002-7316.76.2.229

Eren, M.I., Roos, C.I., Story, B.A., Von Cramon-Taubadel, N., Lycett, S.J., 2014. The role of raw material differences in stone tool shape variation: an experimental assessment. Journal of Archaeological Science 49, 472–487. https://doi.org/10.1016/j.jas.2014.05.034

Ericsson, K.A., 2004. Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains. Academic Medicine 79.

Ferguson, J.R., 2003. An Experimental Test of the Conservation of Raw Material in Flintknapping Skill Acquisition. Lithic Technology 28, 113–131. https://doi.org/10.1080/01977261.2003.11721007

Field, A.S., 2005. Transformations in dividuality: personhood and palaeoliths in the Middle Pleistocene, in: Gamble, C., Porr, M. (Eds.), The Hominid Individual in Context: Archaeological Investigations of Lower and Middle Palaeolithic Landscapes, Locales and Artefacts. Routledge, London, pp. 29–49.

Field, A.S., 2002. The Middle Pleistocene in Transition: Lithic assemblages and changing social relations between OIS 12 and 6 in Europe and Africa. University of Southampton.

Finlay, N., 2015. Kid-knapped Knowledge: Changing Perspectives on the Child in Lithic Studies. Childhood in the Past 8, 104–112. https://doi.org/10.1179/1758571615Z.00000000032

Finlay, N., 2011. Skill and the Substance of Stone. Lithic Technology 36, 189–200. https://doi.org/10.1179/lit.2011.36.2.189

Finlayson, B., Warren, G., 2017. The diversity of hunter-gatherer pasts: an introduction, in: Finlayson, B., Warren, G. (Eds.), The Diversity of Hunter Gatherer Pasts. Oxbow Books, Oxford, pp. 1–14.

Foulds, F., 2014. Invisible Individuals, Visible Groups: On the evidence for individuals and groups at the Lower Palaeolithic site of Caddington, Bedfordshire., in: Foulds, F., W.F., Drinkhall, H.C., Perri, A.R., Clinnick, D.T.G., Walker, J.W.P. (Eds.), Wild Things: Recent Advances in Palaeolithic and Mesolithic Research. Oxbow Books, Oxford, p. NEED.

Friston, K., 2011. Embodied inference: or "I think therefore I am, if I am what I think"., in: Tschacher, W., Bergomi, C. (Eds.), The Implications of Embodiment (Cognition and Communication). Imprint Academic, Exeter, pp. 89–125.

Friston, K., 2010. The free-energy principle: a unified brain theory? Nature Reviews Neuroscience 11, 127. https://doi.org/10.1038/nrn2787

Friston, K., 2009. The free-energy principle: a rough guide to the brain? Trends in cognitive sciences 13, 293–301. https://doi.org/10.1016/j.tics.2009.04.005
Friston, K., Daunizeau, J., Kiebel, S., 2009. Reinforcement Learning or Active Inference? PLOS ONE.
Friston, K., Mattout, J., Kilner, J., 2011. Action understanding and active inference. Biol Cybern 104, 137–160. https://doi.org/10.1007/s00422-011-0424-z
Friston, K.J., Daunizeau, J., Kilner, J., Kiebel, S.J., 2010. Action and behavior: a free-energy formulation.(Report). Biological Cybernetics 102, 227. https://doi.org/10.1007/s00422-010-0364-z
Gabora, L., 2004. Ideas are not replicators but minds are. Biology & Philosophy 19, 127–143.
Gamble, C., Porr, M., 2005. From empty spaces to lived lives: exploring the individual in the Palaeolithic, in: Gamble, C., Porr, M. (Eds.), The Hominid Individual in Context: Archaeological Investigations of Lower and Middle Palaeolithic Landscapes, Locales and Artefacts. Routledge, London, pp. 1–12.
García-Medrano, P., Ashton, N., Moncel, M.-H., Ollé, A., 2020. The WEAP Method: a New Age in the Analysis of the Acheulean Handaxes. Journal of Paleolithic Archaeology. https://doi.org/10.1007/s41982-020-00054-5
García-Medrano, P., Ollé, A., Ashton, N., Roberts, M., 2019. The Mental Template in Handaxe Manufacture: New Insights into Acheulean Lithic Technological Behavior at Boxgrove, Sussex, UK. J Archaeol Method Theory 26, 396–422. https://doi.org/10.1007/s10816-018-9376-0
Gärdenfors, P., 2017. Demonstration and Pantomime in the Evolution of Teaching. Frontiers in Psychology 8, 1–12.
Gärdenfors, P., 2004. Cooperation and the evolution of symbolic communication., in: Kimborough Oller, D., Griebel, U. (Eds.), The Evolution of Communication Systems. MIT Press, Massachusetts, pp. 237–256.
Geribàs, N., Mosquera, M., Vergès, J.-M., 2010. What novice knappers have to learn to become expert stone toolmakers. Journal of Archaeological Science 37, 2857–2870.
Gershman, S., J., Daw, N., D., 2012. Perception, Action, and Utility: The Tangled Skein, in: Rabinovich, M.I., Friston, K.J., Varona, P. (Eds.), Principles of Brain Dynamics: Global State Interactions. MIT Press, Massachusetts, pp. 293–312.
Gibson, J.J., 1979. The ecological approach to visual perception. Houghton Mifflin, Boston.
Giddens, A., 1984. The constitution of society : outline of the theory of structuration. Polity, Cambridge.
Gladwell, M., 2008. Outliers: The story of success. Little, Brown.
Goren-Inbar, N., Alperson, N., Kislev, M.E., Simchoni, O., Melamed, Y., Ben-Nun, A., Werker, E., 2004. Evidence of hominin control of fire at Gesher Benot Ya'aqov, Israel. Science (New York, N.Y.) 304, 725. https://doi.org/10.1126/science.1095443
Goren-Inbar, N., Lister, A., Werker, E., Chech, M., 1994. A butchered elephant skull and associated artifacts from the Acheulian site of Gesher Benot Ya'Aqov, Israel. Paléorient 20, 99. https://doi.org/10.3406/paleo.1994.4604
Gowlett, J., Gamble, C., Dunbar, R., 2012. Human Evolution and the Archaeology of the Social Brain. Current Anthropology 53, 693–722. https://doi.org/10.1086/667994
Gowlett, J.A.J., 2011. The empire of the Acheulean strikes back, in: Sept, J., Pilbeam, D.R. (Eds.), Casting the Net Wide. Papers in Honor of Glynn Isaac and His Approach to Human Origins Research. Oxbow Books, Oxford, pp. 93–114.

Gowlett, J.A.J., 2006. The elements of design form in Acheulian bifaces: modes, modalities, rules and language., in: Goren-Inbar, N., Sharon, G. (Eds.), Axe Age: Acheulian Tool-Making from Quarry to Discard. Equinox, London, pp. 203–221.
Gowlett, J.A.J., 1984. Mental abilities of early man: a look at some hard evidence., in: Foley, R. (Ed.), Human Evolution and Community Ecology. Academic Press, London, pp. 167–192.
Gowlland, G., 2019. The Sociality of Enskilment. null 84, 508–524. https://doi.org/10.1080/00141844.2018.1455726
Gray, W.D., Lindstedt, J.K., 2017. Plateaus, Dips, and Leaps: Where to Look for Inventions and Discoveries During Skilled Performance. Cognitive Science 41, 1838–1870. https://doi.org/10.1111/cogs.12412
Grove, M., 2016. Population density, mobility, and cultural transmission. Journal of Archaeological Science 74, 75–84.
Guthrie, R.D., 1990. Frozen fauna of the Mammoth Steppe : the story of Blue Babe, Ebook central. University of Chicago Press, Chicago.
Haq, M., 2014. A comparative analysis of qualitative and quantitative research methods and a justification for use of mixed methods in social research. Presented at the Annual PhD Conference, University of Bradford School of Management., Bradford Scholars.
Hardaker, T., Dunn, S., 2005. The Flip test - a new statistical measure for quantifying symmetry in stone tools. Antiquity 79.
Hawkins, J., Blakeslee, S., 2005. On intelligence, 1st Owl Books. ed. Henry Holt and Company, New York.
Hecht, E.E., Gutman, D.A., Khreisheh, N., Taylor, S.V., Kilner, J., Faisal, A.A., Bradley, B.A., Chaminade, T., Stout, D., 2015. Acquisition of Paleolithic toolmaking abilities involves structural remodeling to inferior frontoparietal regions. Brain Structure and Function 220, 2315–2331. https://doi.org/10.1007/s00429-014-0789-6
Henrich, J., 2004. Demography and Cultural Evolution: How Adaptive Cultural Processes can Produce Maladaptive Losses: The Tasmanian Case. American Antiquity 69, 197–214. https://doi.org/10.2307/4128416
Henrich, J., Boyd, R., Richerson, P.J., 2008. Five Misunderstandings About Cultural Evolution. Human Nature 19, 119–137.
Hérisson, D., Brenet, M., Cliquet, D., Moncel, M.-H., Richter, J., Scott, B., Van Baelen, A., Di Modica, K., De Loecker, D., Ashton, N., Bourguignon, L., Delagnes, A., Faivre, J.-P., Folgado-Lopez, M., Locht, J.-L., 2016. The emergence of the Middle Palaeolithic in north-western Europe and its southern fringes. Quaternary international 411, 233–283. https://doi.org/10.1016/j.quaint.2016.02.049
Herzlinger, G., Goren-Inbar, N., Grosman, L., 2017. A new method for 3D geometric morphometric shape analysis: The case study of handaxe knapping skill. Journal of Archaeological Science: Reports 14, 163–173. https://doi.org/10.1016/j.jasrep.2017.05.013
Hirose, N., 2002. An ecological approach to embodiment and cognition. Cognitive systems research 3, 289–299. https://doi.org/10.1016/S1389-0417(02)00044-X
Hodgson, D., 2015. The symmetry of Acheulean handaxes and cognitive evolution. Journal of Archaeological Science: Reports 2, 204–208. https://doi.org/10.1016/j.jasrep.2015.02.002
Hodgson, D., 2009. Evolution of the visual cortex and the emergence of symmetry in the Acheulean techno-complex. Comptes rendus - Palevol 8, 93–97. https://doi.org/10.1016/j.crpv.2008.10.009

Högberg, A., Gärdenfors, P., Larsson, L., 2015. Knowing, Learning and Teaching-How Homo Became Docens. Cambridge Archaeological Journal 25, 847–858.
Holmes, N., Spence, C., 2004. The body schema and multisensory representation(s) of peripersonal space. Cogn Process 5, 94–105. https://doi.org/10.1007/s10339-004-0013-3
Hopkinson, T., 2007. The transition from the Lower to the Middle Palaeolithic in Europe and the incorporation of difference. Antiquity 81, 294–307. https://doi.org/10.1017/S0003598X00095181
Hopkinson, T., Nowell, A., White, M.J., 2013. Life Histories, Metapopulation Ecology, and Innovation in the Acheulian. PaleoAnthropology 61–73. https://doi.org/doi:10.4207/PA.2013.ART80
Hosfield, R., 2016. Walking in a Winter Wonderland? Strategies for Early and Middle Pleistocene Survival in Midlatitude Europe. Current Anthropology 57, 653–682. https://doi.org/10.1086/688579
Hosfield, R., 2005. Individuals among palimpsest data: fluvial landscapes in Southern England., in: Gamble, C., Porr, M. (Eds.), The Hominid Individual in Context: Archaeological Investigations of Lower and Middle Palaeolithic Landscapes, Locales and Artefacts. Routledge, London, pp. 220–243.
Hosfield, R., Cole, J., Mcnabb, J., 2018. Less of a bird's song than a hard rock ensemble. Evolutionary Anthropology: Issues, News, and Reviews 27, 9–20. https://doi.org/10.1002/evan.21551
Hutchence, L., Debackere, S., 2018. An evaluation of behaviours considered indicative of skill in handaxe manufacture. Lithics 39, 36–51.
Hutchence, L., Scott, C., 2021. Is Acheulean Handaxe Shape the Result of Imposed 'Mental Templates' or Emergent in Manufacture? Dissolving the Dichotomy through Exploring 'Communities of Practice' at Boxgrove, UK. Cambridge Archaeological Journal 31, 675–686. https://doi.org/10.1017/S0959774321000251
Iliopoulos, A., 2019. Material Engagement Theory and its philosophical ties to pragmatism. Phenomenology and the Cognitive Sciences 18, 39–63. https://doi.org/10.1007/s11097-018-9596-5
Ingold, T., 2013. Making : anthropology, archaeology, art and architecture. Routledge, London.
Ingold, T., 2000. The perception of the environment : essays on livelihood, dwelling & skill. Routledge, London.
Iovita, R., Mcpherron, S.P., 2011. The handaxe reloaded: A morphometric reassessment of Acheulian and Middle Paleolithic handaxes. Journal of Human Evolution 61, 61–74. https://doi.org/10.1016/j.jhevol.2011.02.007
Iovita, R., Tuvi-Arad, I., Moncel, M.-H., Despriée, J., Voinchet, P., Bahain, J.-J., 2017. High handaxe symmetry at the beginning of the European Acheulian: The data from la Noira (France) in context. PLoS ONE 12, e0177063. https://doi.org/10.1371/journal.pone.0177063
Iriki, A., 2006. The neural origins and implications of imitation, mirror neurons and tool use. Current opinion in neurobiology 16, 660–667. https://doi.org/10.1016/j.conb.2006.10.008
Isaac, G., 1977. Olorgesailie: The Archaeology of a Middle Pleistocene Lake Basin in Kenya. Chicago University Press, Chicago.
Keeley, L.H., 1977. An experimental study of microwear traces on selected British Palaeolithic implements. Thesis (D.Phil.)--University of Oxford, 1977.

Keller, C.M., Keller, J.D., 1996. Cognition and tool use : the blacksmith at work, Learning in doing. Cambridge University Press, Cambridge.
Kempe, M., Lycett, S., Mesoudi, A., 2012. An Experimental Test of the Accumulated Copying Error Model of Cultural Mutation for Acheulean Handaxe Size (Acheulean Handaxe Size and Cultural Mutation) 7, e48333. https://doi.org/10.1371/journal.pone.0048333
Key, A., Lycett, S., 2017. Influence of Handaxe Size and Shape on Cutting Efficiency: A Large-Scale Experiment and Morphometric Analysis. J Archaeol Method Theory 24, 514–541. https://doi.org/10.1007/s10816-016-9276-0
Key, A.J.M., 2019. Handaxe shape variation in a relative context. Comptes rendus - Palevol. https://doi.org/10.1016/j.crpv.2019.04.008
Khreisheh, N., 2013. The acquisition of skill in early flaked stone technologies: an experimental study.
Khreisheh, N.N., Davies, D., Bradley, B.A., 2013. Extending Experimental Control 1, 38–46. https://doi.org/10.7183/2326-3768.1.1.37
Klein, R.G., 2009. The human career : human biological and cultural origins, 3rd ed. University of Chicago Press, Chicago, Ill.
Klein, R.G., 2000. Archeology and the evolution of human behavior. Evolutionary Anthropology: Issues, News, and Reviews 9, 17–36. https://doi.org/10.1002/(SICI)1520-6505(2000)9:1<17::AID-EVAN3>3.0.CO
Kline, M.A., 2015. How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. Behavioral and Brain Sciences 38, 1–71.
Knappett, C., 2004. The Affordances of Things: a Post-Gibsonian Perspective on the Relationality of Mind and Matter., in: DeMarrais, E., Gosden, C., Renfrew, C. (Eds.), Rethinking Materiality: The Engagement of Mind with the Material World. McDonald Institute for Archaeological Research, Cambridge, pp. 43–52.
Kohn, M., Mithen, S., 1999. Handaxes: products of sexual selection? Antiquity 73, 518–526.
Lamotte, A., 2012. Le Paléolithique inférieur et moyen des bassins-versants de la Somme et de la Saône: aspects cognitifs de l'environnement pétrographique et considérations techno-typologiques. Habilitation à diriger des recherches dissertation. Université de Lille.
Lamotte, A., 2011. Le Paléolithique inférieur, Le Paléolithique moyen et les complexes de transition avec le Paléolithique supérieur en Haute-Saône : bilan de 7 années de recherches (rapport CIRA DRAC Franche-Comté).
Lamotte, A., 2001. Les industries à bifaces de l'Europe du nord-ouest au pléistocène moyen : l'apport des donnés des gisements du bassin de la Somme, de l'Escaut et de la Baie de St-Brieuc, BAR international series. Archaeopress, Oxford.
Lamotte, A., 1992. Le gisement paléolithique moyen de faciès cambrésien de Gouzeaucourt (Nord). Essai de reconstitution des chaînes opératoires. Bulletin de la Société préhistorique française 89, 37–41. https://doi.org/10.3406/bspf.1992.9494
Lave, J., Wenger, E., 1991. Situated learning : legitimate peripheral participation, Learning in doing. Cambridge University Press, Cambridge.
Lepre, C.J., Roche, H., Kent, D.V., Harmand, S., Quinn, Brugal, J.-P., Texier, P.-J., Lenoble, A., Feibel, C.S., 2011. An earlier origin for the Acheulian. Nature 477, 82. https://doi.org/10.1038/nature10372
Leroyer, M., 2016. Palethnologie acheuléenne : de la technologie bifaciale à l'organisation de la subsistance collective. Etude du site de Boxgrove Eartham pit

(West Sussex, Angleterre) et de deux sites du cours moyen de la Seine. Université Paris 1 Panthéon-Sorbonne.
Li, H., Kuman, K., Lotter, M.G., Leader, G.M., Gibbon, R.J., 2017. The Victoria West: earliest prepared core technology in the Acheulean at Canteen Kopje and implications for the cognitive evolution of early hominids. Royal Society Open Science 4. https://doi.org/10.1098/rsos.170288
Lycett, S.J., 2019. Cultural Transmission from the Last Common Ancestor to the Levallois Reducers, in: Overmann, K.A., Coolidge, F.L. (Eds.), Squeezing Minds From Stones: Cognitive Archaeology and the Evolution of the Human Mind: Cognitive Archaeology and the Evolution of the Human Mind. Oxford University Press, New York, pp. 251–277.
Lycett, S.J., 2015. Cultural evolutionary approaches to artifact variation over time and space: basis, progress, and prospects. Journal of Archaeological Science 56, 21–31.
Lycett, S.J., Schillinger, K., Eren, M.I., Von Cramon-Taubadel, N., Mesoudi, A., 2016. Factors affecting Acheulean handaxe variation: Experimental insights, microevolutionary processes, and macroevolutionary outcomes. Quaternary International 411, 386–401. https://doi.org/10.1016/j.quaint.2015.08.021
Lycett, S.J., Von Cramon-Taubadel, N., 2008. Acheulean variability and hominin dispersals: a model-bound approach. Journal of archaeological science 35, 553–562. https://doi.org/10.1016/j.jas.2007.05.003
Lycett, S.J., Von Cramon-Taubadel, N., Foley, R.A., 2006. A crossbeam co-ordinate caliper for the morphometric analysis of lithic nuclei: a description, test and empirical examples of application. Journal of archaeological science 33, 847–861. https://doi.org/10.1016/j.jas.2005.10.014
MacDonald, K., 2007. Ecological hypotheses for human brain evolution: evidence for skill and learning processes in the ethnographic literature on hunting., in: Roebroeks, W. (Ed.), Guts and Brains: An Integrative Approach to the Hominin Record. Leiden University Press, Leiden, pp. 107–132.
Machin, A., 2009. The role of the individual agent in Acheulean biface variability. Journal of Social Archaeology 9, 35–58.
Machin, A.J., Hosfield, R.T., Mithen, S.J., 2007. Why are some handaxes symmetrical? Testing the influence of handaxe morphology on butchery effectiveness. Journal of Archaeological Science 34, 883–893. https://doi.org/10.1016/j.jas.2006.09.008
Maier, A., 2015. The Central European Magdalenian: Regional Diversity and Internal Variability. Springer, Dordecht.
Malafouris, L., 2020a. How does thinking relate to tool making? Adaptive Behavior 1–15. https://doi.org/10.1177/1059712320950539
Malafouris, L., 2020b. Thinking as "Thinging": Psychology With Things. Current directions in psychological science : a journal of the American Psychological Society 29, 3–8. https://doi.org/10.1177/0963721419873349
Malafouris, L., 2013. How Things Shape the Mind [electronic resource] : A Theory of Material Engagement, Book collections on Project MUSE. MIT Press, Cambridge, Massachusetts.
Malafouris, L., 2010. Knapping Intentions and the Marks of the Mental., in: Malafouris, L., Renfrew, C. (Eds.), The Cognitive Life of Things: Recasting the Boundaries of the Mind. McDonald Institute for Archaeological Research, Cambridge, pp. 13–22.

Malafouris, L., 2008. Beads for a Plastic Mind: the 'Blind Man's Stick' (BMS) Hypothesis and the Active Nature of Material Culture. CAJ 18, 401–414. https://doi.org/10.1017/S0959774308000449

Malinsky-Buller, A., 2016. Lost and found: Technological trajectories within Lower/Middle Paleolithic transition in Western Europe, North of the Pyrenees. Quaternary International 409, 104–148. https://doi.org/10.1016/j.quaint.2015.09.079

Maravita, A., Iriki, A., 2004. Tools for the body (schema). Trends in cognitive sciences 8, 79–86. https://doi.org/10.1016/j.tics.2003.12.008

Marchand, T., 2007. Crafting Knowledge: The Role of "Parsing and Production" in the Communication of Skill-based Knowlede among Masons., in: Harris, M. (Ed.), Ways of Knowing: Anthropological Approaches to Crafting Experience and Knowledge. Berghahn, Oxford, pp. 181–202.

Marshall, G.D., Gamble, C.G., Roe, D.A., Dupplaw, D., 2002. Acheulian biface database. ADS, York.

Maturana, H.R., Varela, F.J., 1972. Autopoiesis and cognition : the realization of the living. Reidel, Dordrecht.

McNabb, J., 2017. Journeys in space and time. Assessing the link between Acheulean handaxes and genetic explanations. Journal of Archaeological Science: Reports 13, 403–414. https://doi.org/10.1016/j.jasrep.2017.04.010

McNabb, J., Cole, J., 2015. The mirror cracked: Symmetry and refinement in the Acheulean handaxe. Journal of Archaeological Science: Reports 3, 100–111. https://doi.org/10.1016/j.jasrep.2015.06.004

McNabb, J., Cole, J., Hoggard, C.S., 2018. From side to side: Symmetry in handaxes in the British Lower and Middle Palaeolithic. Journal of Archaeological Science: Reports 17, 293–310. https://doi.org/10.1016/j.jasrep.2017.11.008

McPherron, S., 1995. A Reexamination of British Biface Data. Lithics 16.

McPherron, S.P., 2000. Handaxes as a Measure of the Mental Capabilities of Early Hominids. Journal of Archaeological Science 27, 655–663. https://doi.org/10.1006/jasc.1999.0467

McPherron, S.P., 1999. Ovate and Pointed Handaxe Assemblages: Two points make a line. Préhistoire Européenne 14, 9–32.

McPherron, S.P., 1994. A Reduction Model for Variability in Acheulean Biface Morphology. University of Pennsylvania, Pennsylvania.

Milks, A., Parker, D., Pope, M., 2019. External ballistics of Pleistocene hand-thrown spears: experimental performance data and implications for human evolution. Scientific Reports 9, 820. https://doi.org/10.1038/s41598-018-37904-w

Miller, L.E., Longo, M.R., Saygin, A.P., 2014. Tool morphology constrains the effects of tool use on body representations.(Report)(Author abstract). Journal of Experimental Psychology: Human Perception and Performance 40, 2143. https://doi.org/10.1037/a0037777

Mitchell, J.C., 1998. A use-wear analysis of selected British lower palaeolithic handaxes with special reference to the site of Boxgrove (West Sussex) : a study incorporating optical microscopy, computer aided image analysis and experimental archaeology. Thesis (D.Phil.)--University of Oxford, 1998.

Morgan, T.J.H., Uomini, N.T., Rendell, L.E., Chouinard-Thuly, L., Street, S.E., Lewis, H.M., Cross, C.P., Evans, C., Kearney, R., de La Torre, I., Whiten, A., Laland, K.N., 2015. Experimental evidence for the co-evolution of hominin tool-making teaching and language. Nature Communications 6. https://doi.org/10.1038/ncomms7029

Muller, A., Clarkson, C., Shipton, C., 2017. Measuring behavioural and cognitive complexity in lithic technology throughout human evolution. Journal of Anthropological Archaeology 48, 166–180.

Newcomer, M.H., 1971. Some quantitative experiments in handaxe manufacture. World Archaeology 3, 85–94. https://doi.org/10.1080/00438243.1971.9979493

Nonaka, T., Bril, B., Rein, R., 2010. How do stone knappers predict and control the outcome of flaking? Implications for understanding early stone tool technology. Journal of Human Evolution 59, 155–167.

O'Connor, E., 2006. Glassblowing Tools: Extending the Body Towards Practical Knowledge and Informing a Social World. Qual Sociol 29, 177–193. https://doi.org/10.1007/s11133-006-9013-1

O'Regan, J.K., Noë, A., 2001. A sensorimotor account of vision and visual consciousness. Behav. Brain Sci. 24, 939–973. https://doi.org/10.1017/S0140525X01000115

Pargeter, J., Khreisheh, N., Stout, D., 2019. Understanding stone tool-making skill acquisition: Experimental methods and evolutionary implications. Journal of Human Evolution 133, 146–166. https://doi.org/10.1016/j.jhevol.2019.05.010

Pargeter, J., Kreisheh, N., Shea, J.J., Stout, D., 2020. Knowledge vs. know-how? Dissecting the foundations of stone knapping skill. Journal of Human Evolution 145, 102807. https://doi.org/10.1016/j.jhevol.2020.102807

Pelegrin, J., 2005. Remarks about Archaeological Techniques and Methods of Knapping: Elements of a Cognitive Approach to Stone Knapping, in: Roux, V., Bril, B. (Eds.), Stone Knapping: A Uniquely Hominin Behaviour. McDonald Institute for Archaeological Research, Cambridge, pp. 23–34.

Pelegrin, J., 1993. A framework for analysing prehistoric stone tools manufacture and a tentative application to some early lithic industries., in: Berthelet, A., Chavaillon, J. (Eds.), The Use of Tools by Human and Non-Human Primates: A Symposium of the Fyssen Foundation. Clarendon Press, Oxford, pp. 302–314.

Pelegrin, J., 1990. Prehistoric lithic technology: some aspects of research. Archaeological Review of Cambridge 9, 116–125.

Pigeot, N., 1990. Technical and Social Actors: Flintknapping Specialists and Apprentices at Magdalenian Etiolles. Archaeological Review from Cambridge 9, 126–141.

Ploux, S., 1989. Approche archéologique de la variabilité des comportements individuels. L'exemple de quelques tailleurs magdaléniens à Pincevent. University Paris X, - Nanterre, Nanterre.

Polanyi, M., 1958. Personal knowledge : towards a post-critical philosophy. Routledge & Kegan Paul, London.

Pope, M., 2002. The significance of biface-rich assemblages: An examination of behavioural controls on lithic assemblage formation in the Lower Palaeolithic. University of Southampton.

Pope, M., Parfitt, S., Roberts, M., 2020a. The Horse Butchery Site : A high-resolution record of Lower Palaeolithic hominin behaviour at Boxgrove, UK. SpoilHeap Publications, Portslade.

Pope, M., Parfitt, S., Roberts, M., 2020b. The Horse Butchery Site : A high-resolution record of Lower Palaeolithic hominin behaviour at Boxgrove, UK. SpoilHeap Publications, Portslade.

Pope, M., Russel, K., Watson, K., 2006. Biface Form and Structured Behaviour in the Acheulean. Lithics 27, 44–57.

Porr, B., Wöörgöötter, F., 2003. Isotropic Sequence Order Learning. Neural Computation 15, 831–864. https://doi.org/10.1162/08997660360581921

Powell, A., Shennan, S., Thomas, M., 2009. Late Pleistocene demography and the appearance of modern human behaviour. Science 324, 1298–1301.
Rabinovich, R., Biton, R., 2011. The Early–Middle Pleistocene faunal assemblages of Gesher Benot Ya'aqov: Inter-site variability. Journal of Human Evolution 60, 357–374. https://doi.org/10.1016/j.jhevol.2010.12.002
Ramsey, G., De Block, A., 2015. Is Cultural Fitness Hopelessly Confused? The British Journal for the Philosophy of Science 0, 1–24.
Rietveld, E., Kiverstein, J., 2014. A Rich Landscape of Affordances. Ecological psychology 26, 325–352. https://doi.org/10.1080/10407413.2014.958035
Rightmire, G.P., 2008. Homo in the middle pleistocene: Hypodigms, variation, and species recognition. Evolutionary Anthropology: Issues, News, and Reviews 17, 8–21. https://doi.org/10.1002/evan.20160
Rightmire, G.P., 2004. Brain size and encephalization in early to Mid-Pleistocene Homo. American Journal of Physical Anthropology 124, 109–123. https://doi.org/10.1002/ajpa.10346
Roberts, M.B., Parfitt, S.A., 1999. Boxgrove: A Middle Pleistocene hominid site at Eartham Quarry, Boxgrove, West Sussex. English Heritage, London.
Roberts, M.B., Parfitt, S.A., Pope, M.I., Wenban-Smith, F.F., Macphail, R.I., Locker, A., Stewart, J.R., 1997. Boxgrove, West Sussex: Rescue Excavations of a Lower Palaeolithic Landsurface (Boxgrove Project B, 1989–91). Proc. Prehist. Soc. 63, 303–358. https://doi.org/10.1017/S0079497X00002474
Roe, D., 1980. Introduction: Precise moments in remote time. null 12, 107–108. https://doi.org/10.1080/00438243.1980.9979786
Roe, D., 1968. British Lower and Middle Palaeolithic handaxe groups. Proceedings of the Prehistoric Society 34, 1–82.
Rose, M., 2004. The Mind at Work: Valuing the intelligence of the American worker. Penguin, London.
Roux, Valentine., Corbetta, Danièla., 1989. The potter's wheel : craft specialization and technical competence. Oxford & IBH Pub. Co, New Delhi.
Sackett, J., 2014. François Bordes and the Old Stone Age. Bulletin of the History of Archaeology 24, Art. 3. https://doi.org/10.5334/bha.243
Schick, K., Toth, N., 2017. Acheulean Industries of the Early and Middle Pleistocene, Middle Awash, Ethiopia. l'anthropologie 121, 451–491.
Schick, K.D., Toth, N.P., 1993. Making silent stones speak : human evolution and the dawn of technology. Weidenfeld and Nicolson, London.
Schillinger, K., Mesoudi, A., Lycett, S.J., 2016. Copying error, evolution, and phylogenetic signal in artifactual traditions: An experimental approach using "model artifacts." Journal of Archaeological Science 70, 23–34. https://doi.org/10.1016/j.jas.2016.04.013
Schillinger, K., Mesoudi, A., Lycett, S.J., 2015. The impact of imitative versus emulative learning mechanisms on artefactual variation: implications for the evolution of material culture. Evolution and Human Behaviour 36, 446–455.
Schillinger, K., Mesoudi, A., Lycett, S.J., 2014. Copying Error and the Cultural Evolution of "Additive" vs. "Reductive" Material Traditions: An Experimental Assessment. American Antiquity 79, 128–143. https://doi.org/10.7183/0002-7316.79.1.128
Scott, B., Ashton, N., 2010. The Early Middle Palaeolithic: The European Context, in: Ashton, N., Lewis, S.G., Stringer, C. (Eds.), The Ancient Human Occupation of Britain, Developments in Quaternary Science. Elsevier, Amsterdam, pp. 91–112.
Sharon, G., 2010. Large flake Acheulian. Quaternary International 223 224, 226.

Shea, J.J., 2020. Prehistoric stone tools of Eastern Africa : a guide, Cambridge core. Cambridge University Press, Cambridge.
Shea, J.J., 2006. Child's Play: Reflections on the Invisibility of Children in the Paleolithic Record. Evolutionary Anthropology 15, 212–216.
Shipton, C., 2019. The Evolution of Social Transmission in the Acheulean, in: Coolidge, F.L., Overmann, K. (Eds.), Squeezing Minds from Stones. Oxford University Press, Online/Web.
Shipton, C., 2018. Biface Knapping Skill in the East African Acheulean: Progressive Trends and Random Walks. African Archaeological Review 35, 107–131. https://doi.org/10.1007/s10437-018-9287-1
Shipton, C., 2010. Imitation and Shared Intentionality in the Acheulean. CAJ 20, 197–210. https://doi.org/10.1017/S0959774310000235
Shipton, C., Clarkson, C., 2015. Handaxe reduction and its influence on shape: An experimental test and archaeological case study. Journal of Archaeological Science: Reports 3 3, 408–419.
Shipton, C., Nielsen, M., 2018. The Acquisition of Biface Knapping Skill in the Acheulean, in: Di Vincenzo, F., Di Paolo, L.D. (Eds.), The Evolution of Primate Social Cognition. Springer, Dordecht.
Shipton, C., White, M., 2020. Handaxe types, colonization waves, and social norms in the British Acheulean. Journal of Archaeological Science: Reports 31, 102352. https://doi.org/10.1016/j.jasrep.2020.102352
Solodenko, N., Zupancich, A., Cesaro, S.N., Marder, O., Lemorini, C., Barkai, R., 2015. Fat Residue and Use-Wear Found on Acheulian Biface and Scraper Associated with Butchered Elephant Remains at the Site of Revadim, Israel.(Research Article)(Report). PLoS ONE 10, e0118572. https://doi.org/10.1371/journal.pone.0118572
Sommé, J., 1977. Les plaines du Nord de la France et leur bordure : étude géomorphologique. Atelier reproduction des thèses, Université de Lille III, Lille : Paris.
Soriano, S., 2000. Outillage bifacial et outillage sur éclat au Paléolithique ancien et moyen: coexistence and interaction. Université Paris X - Nanterre, Paris.
Stapert, D., 2007. Youngsters knapping flint near the campfire: An alternative view of site K at Maastricht-Belvédère (The Netherlands). Archäologisches Korrespondenzblatt 37, 19–35.
Stout, D., 2011. Stone toolmaking and the evolution of human culture and cognition. Philosophical Transactions of The Royal Society B: Biological Sciences 366, 1050–1059.
Stout, D., 2008. Tools and Brain Size: Technology and Human Brain Evolution. General Anthroplogy 15, 1–5.
Stout, D., 2002. Skill and Cognition in Stone Tool Production: An Ethnographic Case Study from Irian Jaya. Current Anthropology 43, 693–722.
Stout, D., Apel, J., Commander, J., Roberts, M.B., 2014. Late Acheulean technology and cognition at Boxgrove, UK. Journal of Archaeological Science 41, 576–590.
Stout, D., Chaminade, T., 2012. Stone tools, language and the brain in human evolution. Philosophical Transactions of The Royal Society B 367, 75–87.
Stout, D., Chaminade, T., 2009. Making Tools and Making Sense: Complex, Intentional Behaviour in Human Evolution. Cambridge Archaeological Journal 19, 85–96.
Stout, D., Hecht, E., Khreisheh, N., Bradley, B., Chaminade, T., 2015. Cognitive Demands of Lower Paleolithic Toolmaking. PLoS ONE 10, e0121804.

Stout, D., Khreisheh, N., 2015. Skill Learning and Human Brain Evolution: An Experimental Approach. Cambridge Archaeological Journal 25, 867–875.

Terashima, H., 2013. The Evolutionary Development of Learning and Teaching Strategies in Human Societies., in: Akazawa, T., Nishiaki, Y., Aoki, K. (Eds.), Dynamics of Learning in Neanderthals and Modern Humans. Springer, London, pp. 141–150.

Texier, P.-J., 2018. Technological Assets for the Emergence of the Acheulean? Reflections on the Kokiselei 4 Lithic Assemblage and Its Place in the Archaeological Context of West Turkana, Kenya, in: Gallotti, R., Mussi, M. (Eds.), The Emergence of the Acheulean in East Africa and Beyond: Contributions in Honor of Jean Chavaillon. Springer International Publishing, Cham, pp. 33–52. https://doi.org/10.1007/978-3-319-75985-2_3

Thieme, H., 1997. Lower Palaeolithic hunting spears from Germany. Nature 385, 807. https://doi.org/10.1038/385807a0

Torres, C., Preysler, J.B., 2020. Experts Also Fail: a New Methodological Approach to Skills Analysis in Lithic Industries. Journal of Paleolithic Archaeology. https://doi.org/10.1007/s41982-020-00063-4

Tuffreau, A., Bouchet, J.-P., 1985. Le gisement acheuléen de la Vallée du Muid à Gouzeaucourt (Nord). Bulletin de la Société préhistorique française 82, 291–306. https://doi.org/10.3406/bspf.1985.8643

Tuffreau, A., Lamotte, A., Goval, É., 2008. Les industries acheuléennes de la France septentrionale. L'Anthropologie 112, 104–139. https://doi.org/10.1016/j.anthro.2008.01.003

Ungar, P., Sponheimer, M., 2011. The Diets of Early Hominins. Science 334, 190–193. https://doi.org/10.1126/science.1207701

Ungar, P.S., Grine, F.E., Teaford, M.F., El Zaatari, S., 2006. Dental microwear and diets of African early Homo. Journal of Human Evolution 50, 78–95. https://doi.org/10.1016/j.jhevol.2005.08.007

Vaesen, K., 2012. The cognitive bases of human tool use. Behav Brain Sci 35, 203–218. https://doi.org/10.1017/S0140525X11001452

Van Arsdale, A.P., 2013. Homo erectus - A Bigger, Smarter, Faster Hominin lineage. Nature Education Knowledge 4, 2.

Walker, A., Leakey, R.E., 1993. The Nariokotome Homo erectus skeleton. Springer-Verlag, Berlin.

Wenban-Smith, F.F., 2004. Handaxe typology and Lower Palaeolithic cultural development: crons, cleavers and two giant handaxes from Cuxton. Lithics 25, 11–21.

Wenban-Smith, F.F., 2001. As represented by the Solent River: handaxes from Highfield, Southampton., in: Wenban-Smith, F.F., Hosfield, R.T. (Eds.), Palaeolithic Archaeology of the Solent River. Lithics Studies Society, London, pp. 57–69.

Wenban-Smith, F.F., 1989. The use of canonical variates for determination of biface manufacturing technology at boxgrove lower palaeolithic site and the behavioural implications of this technology. Journal of Archaeological Science 16, 17–26. https://doi.org/10.1016/0305-4403(89)90053-8

Wenger, E., 1998. Communities of practice : learning, meaning, and identity, Learning in doing. Cambridge University Press, Cambridge.

White, M., 1998. On the Significance of Acheulean Biface Variability in Southern Britain. Proceedings of the Prehistoric Society 64, 15–44.

White, M., Foulds, F., 2018. Symmetry is its own reward: On the character and significance of Acheulean handaxe symmetry in the Middle Pleistocene. Antiquity 92, 304–319. https://doi.org/10.15184/aqy.2018.35

White, M., Scott, B., Ashton, N., 2006. The Early Middle Palaeolithic in Britain: archaeology, settlement history and human behaviour. Journal of Quaternary Science 21, 525–541. https://doi.org/10.1002/jqs.1034

White, M.J., 2006. Axeing Cleavers: reflections on broad-tipped large cutting tools in the British Lower and Middle Palaeolithic., in: Goren-Inbar, N., Sharon, G. (Eds.), Axe Age: Acheulean Toolmaking, from Quarry to Discard. Equinox, London, pp. 365–381.

Whiten, A., 2015. Experimental studies illuminate the cultural transmission of percussive technologies in Homo and Pan. Philosophical Transactions of The Royal Society B 370, 20140359.

Wilkins, J., Chazan, M., 2012. Blade production ~500 thousand years ago at Kathu Pan 1, South Africa: support for a multiple origins hypothesis for early Middle Pleistocene blade technologies. Journal of Archaeological Science 39, 1883–1900. https://doi.org/10.1016/j.jas.2012.01.031

Wilkins, J., Schoville, B.J., Brown, K.S., Chazan, M., 2012. Evidence for early hafted hunting technology. Science (New York, N.Y.) 338, 942. https://doi.org/10.1126/science.1227608

Williams-Hatala, E.M., Hatala, K.G., Key, A., Dunmore, C.J., Kasper, M., Gordon, M., Kivell, T.L., 2020. Kinetics of stone tool production among novice and expert tool makers. American journal of physical anthropology e24159. https://doi.org/10.1002/ajpa.24159

Wilson, F.R., 1999. The hand : how its use shapes the brain, language, and human culture. Vintage, New York.

Winton, V., 2005. An investigation of knapping- skill development in the manufacture of Palaeolithic handaxes., in: Roux, V., Bril, B. (Eds.), Stone Knapping: A Uniquely Hominin Behaviour. McDonald Institute for Archaeological Research, Cambridge, pp. 109–116.

Wrangham, R.W., 2010. Catching fire [electronic resource] : how cooking made us human, Alexander Street anthropology. Basic Books, New York.

Wymer, J., 1968. Lower Palaeolithic Archaeology in Britain as represented by the Thames Valley. John Baker, London.

Wynn, T., 2002. Archaeology and cognitive evolution. Behavioral and Brain Sciences 25, 389–402.

Wynn, T., 1995. Handaxe enigmas. World Archaeology: Symbolic Aspects of Early Technologies 27, 10–24. https://doi.org/10.1080/00438243.1995.9980290

Wynn, T., Berlant, T., 2019. The Handaxe Aesthetic, in: Overmann, K.A., Coolidge, F.L. (Eds.), Squeezing Minds From Stones: Cognitive Archaeology and the Evolution of the Human Mind: Cognitive Archaeology and the Evolution of the Human Mind. Oxford University Press, New York, pp. 278–303.

Wynn, T., Gowlett, J., 2018. The handaxe reconsidered. Evolutionary Anthropology: Issues, News, and Reviews 27, 21–29. https://doi.org/10.1002/evan.21552

Appendix.*

This appendix provides a detailed explanation of Linear Discriminant Analysis (LDA) that I have used throughout this thesis. LDA was carried out on quantitative data recorded from both experimental and archaeological handaxe assemblages. The data used is outlined in my Methodology (Chapter 6) and the analysis was performed in PAST version 4.0 software.

LDA is a widely used type of Discriminant Function Analysis (DFA) and calculates the probability of group membership based on predictor variables that have been calibrated on a set of measured variables with known grouping labels. Like PCA, Linear Discriminant Analysis can be used as a tool for reducing the dimensionality of data, but, unlike PCA, it assumes separate distributions for the different groups present in the data.

A critical part of the LDA assumption is that the same dimensions for distinct groups have identical covariance. However, if individual groups are not identically distributed a separation can still be performed if the measured values for the different groups diverge enough (i.e., if the means of the different groups vary by more than both groups' standard deviation/variance).

LDA utilises these distributions to determine a coordinate transformation from the measured variables to a new and ordered set of "canonical coordinates" for which the overlap between the multiple category distributions is minimal for the first coordinate (similarly to a PCA analysis). The dimensional reduction enables the separation between the categories in the two-dimensional space consisting of the two first canonical coordinates, which is shown in Figure 1 below. The coordinate transformation to this

optimal two-dimensional subspace also indicates what criteria best predict inclusion in a group.

Each axis is a linear combination of the original data coordinate axes, that is, the value along each coordinate axis gets multiplied by a value known either as the slope, coefficient, or weight, and these products get summed to give the value along the new axis. The new coordinate axes each explain a percentage of the variance in the data sample, much like a PCA.

The loading score for each criterion on an axis relates to the contribution it makes to the axis (direction of the axis). The loading score for criteria can be either positive or negative but the largest loadings are the ones that dominate the axis direction. In the example Table 1 outlined below, Axis 1, which explains 81.99% of variance, is dominated by Edge Straightness which has the highest loading score – this would have been equally the case if Edge Straightness had a similarly strong negative loading value.

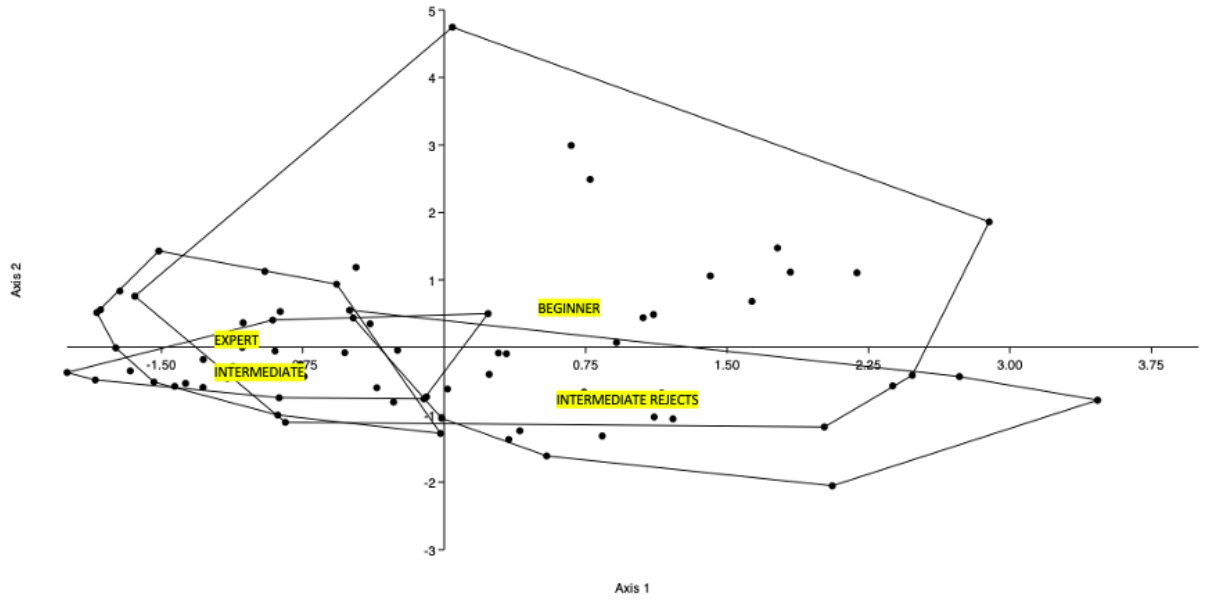
	Axis 1 (81.99%)	**Axis 2 (14.57%)**	**Axis 3 (3.44%)**
Refinement	-2.4187	-9.7687	-0.65058
Th/L	6.4948	20.035	-9.8561
Prof Asymm	0.84593	4.7799	3.1278
Plan ASymm	8.7605	-5.0676	9.8187
Edge Straight	23.147	-5.0886	26.05

Appendix Table 1. Example loading scores for criteria used in LDA of experimental handaxe assemblage.

As outlined in Table 2 below, these loading values translate to 'scores' for each individual handaxe on each axis. These scores can then be plotted similarly to that done with a PCA (see Figure 1).

	Axis 1	**Axis 2**	**Axis 3**
KL_1.1	-1.8424	-0.25096	0.036316
KL_1.2	-1.6024	-1.2764	-0.086311
KL_1.3	-1.1011	0.70396	0.2573
KL_LIV.6	0.067839	-0.16486	-3.1333
KL_LIV.7	-1.1294	-0.39115	0.26976
KL_LIV.8	0.043899	1.0816	1.9264
ADAMTURN.3	1.9378	0.79219	-0.38057
ADAMTURN.4	0.66273	0.65365	0.029689
ADAMTURN.5	0.11844	0.61509	0.38562
LH_KL.1	0.40395	-1.2779	0.66086
LH_KL.2	-0.70899	-0.911	-0.24377
CS3.5	0.75181	0.31157	0.62873
CS3.6	0.33314	2.3676	0.92219
CS3.7	-0.60698	-1.7599	-1.0494

Appendix Table 2. A sample of LDA scores for handaxes from my experimental assemblage.



Appendix Figure 1. Example LDA scatter plot.

With these scores and the known group labels, LDA derives a linear function that separates the specific groups along the individual canonical coordinate axes. With this classification criterion, LDA allocates samples (i.e., bifaces) to the different groups and quantifies how successful the classification function is. It does this by listing the posterior probabilities of each sample/biface to each group. Posterior probabilities measure the strength of the classification, as well as the percentage certainty with which the classification was made. Classifications can then be presented in a table (see Table 3) where each sample's group classification is given alongside its original, known grouping.

Point	**Given group**	**Classification**
KL_1.1	Expert	Expert
KL_1.2	Expert	Expert

KL_1.3	Expert	Intermediate
CS_LIV1.6	Intermediate	Intermediate
CS_LIV1.7	Intermediate	Intermediate
CS_LIV1.8	Intermediate	Expert
ADAMTURN.3	Novice	Novice
ADAMTURN.4	Novice	Intermediate Draft
LH_KL.1	Novice	Novice
LH_KL.2	Novice	Expert
CS3.5	Intermediate Draft	Intermediate Draft
CS3.6	Intermediate Draft	Intermediate Draft
CS3.7	Intermediate Draft	Expert

Appendix Table 3. Sorting table illustrating the classification of a sample of experimental handaxe specimens in comparison to their original, given grouping.

Comparing the accuracy of the classification to its original given group generates the predictive accuracy of your variables for classifying groups in a dataset. This can be presented in a table (see Table 4) and as a percentage (i.e., Table 4 shows that 84 out of 130 samples/bifaces were predicted correctly, which results in a predictive accuracy percentage of 64.62%).

	Expert	**Intermediate**	**Novice**	**Intermediate Draft**	**Total**
Expert	24	9	0	5	38
Intermediate	8	19	1	1	29
Novice	3	3	25	7	38

Intermediate Draft	3	3	3	16	25
Total	38	34	29	29	130

Appendix Table 4. Table showing predictive accuracy of LDA by group.

However, using the entire dataset measures the ‘re-substitution error' - how well the samples are classified when all the samples are used to generate an axis. The issue with using all the samples in a dataset is that every sample is used not only to build the axis, but to test its predictive accuracy as well. To better evaluate the LDA, one should carry out ‘jackknifing’, or leave-one-out cross-validation, instead. This excludes one sample/biface, generates an axis using the existing data (minus that individual sample), and uses that axis to classify the excluded biface sample. The advantage of this ‘jackknifing' method is that it is not self-fulfilling and does not overestimate the predictive power of the LDA performed. If we take the example from Table 4 and ‘jackknife’ the results, the predictive accuracy percentage of the same dataset is reduced to 54.62% as can be seen in Table 5 with the total of correctly predicted samples being reduced from 84 to 71 out of 130 samples/bifaces. Moreover, in Table 6, ‘jackknifing’ shows whether the original classification was correct or not without using itself in the data sample that predicts its classification. As you can see, ‘jackknifing’ can confirm, in the case of handaxe ‘KL_1.3 ’ that the original classification as intermediate was robust, whereas, in the case of ‘CS3.5’, ‘jackknifing' overturned the original classification, indicating the first result could stem from including the sample when training the classifier.

	Expert	Intermediate	Novice	Intermediate Draft	Total
Expert	21	11	1	5	38
Intermediate	11	15	1	2	29
Novice	3	4	22	9	38
Intermediate Draft	4	3	5	13	25
Total	39	33	29	29	130

Appendix Table 5. Table showing predictive accuracy of 'jackknifed' LDA by group.

Point	**Original skill-level category**	**Classification**	**Jackknifed**
KL_1.1	Expert	Expert	Expert
KL_1.2	Expert	Expert	Expert
KL_1.3	Expert	Intermediate	Intermediate
CS_LIV1.6	Intermediate	Intermediate	Intermediate
CS_LIV1.7	Intermediate	Intermediate	Intermediate
CS_LIV1.8	Intermediate	Expert	Expert
ADAMTURN.3	Novice	Novice	Novice
ADAMTURN.4	Novice	Intermediate Draft	Intermediate Draft
LH_KL.1	Novice	Novice	Novice
LH_KL.2	Novice	Expert	Expert
CS3.5	Intermediate Draft	Intermediate Draft	Novice
CS3.6	Intermediate Draft	Intermediate Draft	Intermediate Draft
CS3.7	Intermediate Draft	Expert	Expert

Appendix Table 6. Sorting table illustrating the classification and ‘jackknifed’ classification of a sample of experimental handaxe specimens in comparison to their original, given grouping.

Where the group (i.e., skill-level category) a biface belongs to is priorly known, as in the experimental assemblage, the above process/method is carried out in this thesis. For each LDA conducted I will report the following:

- A table of loadings – highlighting the variables that contribute to each discriminant function.
- The percentage of variation within each group explained by each discriminant function.
- A plot of the LDA classification, showing the group(s)/specimen(s) plotted against the LDA function(s)/axis(es).
- The jackknifed accuracy.

When the skill-level category a biface belongs to is undetermined prior to investigation, such as in the investigation of the archaeological assemblage from Gouzeaucourt level G, the LDA conducted on the experimental assemblage will generate a classification of ‘knowns' that can be applied to the archaeological data. The experimental dataset has practically equal numbers of samples for each of the skill levels, hence, it can be taken as representative of the underlying distribution of the different measured values for all skill levels. With respect to the LDA presented in this thesis, these 'knowns’ relate to criteria thresholds that reflect different skill-level categories (i.e., Expert, Intermediate, Novice). Therefore, when applied to the archaeological data, the ‘known’ classification of what represents ‘Expert’, for example, will predict the groups within the undetermined dataset that fall within the classification range of ‘Expert' as developed by the experimental assemblage LDA.

To achieve this, the measurements of the archaeological data need to be standardised and normalised to the distribution of all skill levels, estimated from the experimental data. This ensures that the scores for each axis, as seen in Tables 1 and 2 above, are determined consistently, i.e., Experts produce samples/bifaces that are similar outliers with respect to the total distribution of skill levels. For each criterion used in the Experimental assemblage LDA one had to correct the archaeological dataset with the mean and the standard deviation of the experimental dataset – e.g., correcting archaeological Edge Straightness data with the mean and standard deviation of that of the experimental dataset and repeating for each individual criterion until you are done.

Classifying the archaeological bifaces by skill-level category involved determining the coordinates of each sample/biface with respect to the canonical axes found through the LDA. Mathematically, this corresponds to projecting the measured values for each sample/biface onto the canonical axes. This is achieved by multiplying each measured criterion of a sample/biface by its normalised loading along the desired canonical axis. Since the axis loadings from the LDA account for the variance along each axis, the axes first need to be normalised to unit length, i.e., 1, such that only the direction of the axis remains. If experimental loadings were used to assess the GZnivG assemblage without normalising them, the archaeological data would have been affected by introducing dimensional bias inherent in the experimental data. Removing the length of the experimental loading axes and keeping only their direction (i.e., variance), allowed the archaeological data to be assessed fairly by transposing it directly onto the frame of variance exhibited by the known experimental data. Therefore, normalisation of the

experimental loadings allowed both experimental and archaeological datasets to be compared within the same frame of reference.

Normalising the loadings for each axis then involves dividing each loading in Table 1 by the length of the axis in question. The length of each axis can be determined from Pythagoras' theorem in higher dimensions, i.e., the length of axis 1 in Table 1 is equal to the square root of the sum of the squares of its loadings:

$$L_1 = \sqrt{(refinement_1)^2 + (Th/L_1)^2 + (ProfAsymm_1)^2 + (PlanAsymm_1)^2 + (EdgeStraight_1)^2}$$

Subsequently, the measured values for each sample/biface are multiplied by their normalised loadings and then summed together to give the desired axis score (see Catalogue 3, Sheet 3.3).

Provided with the axis scores along the two canonical coordinate axes, both the archaeological and experimental data can be plotted in the manner shown in Figure 1. The same linear separation criteria used to analyse experimental data can be used to classify the archaeological samples/bifaces according to the ability of the knapper who made them.

Appendix References

Davis, J.C., 2002. Statistics and data analysis in geology, 3rd ed. John Wiley and Sons, New York.

Holland, S., 2021. Discriminant Function Analysis. Data Analysis in the Geosciences GEOL 8370. URL http://strata.uga.edu/8370/lecturenotes/discriminantFunctionAnalysis.html (accessed 1.21.21).

Maindonald, J.H., Braun, J., 2003. Data analysis and graphics using R : an example-based approach, Cambridge series on statistical and probabilistic mathematics. Cambridge University Press, Cambridge.