Subjectivity Inherent In By-Eye Symmetry Judgements and the Large Cutting Tools at the Cave of Hearths, Limpopo Province, South Africa

Dave Underhill
Centre for the Archaeology of Human Origins, University of Southampton

The Stone Age of South Africa is an area of study due for a renaissance, and there is a real need for unification of the extant evidence. As a beginning to this, new methodologies have been proposed. This paper tackles the issue of symmetry, specifically the subjectivity involved in by-eye judgements. Assumptions of subjectivity, however, are not proof: presented here is a critical analysis of the inherent bias of by-eye symmetry judgements. Ultimately it is clear that the method contains a level of subjectivity which strips it of any analytical value. The by-eye judgement of symmetry is replaced by the more robust Flip Test computer program, and a brief study is made of the Large Cutting Tools (LCT) at a vitally important, yet often overlooked, site dating from the Pleistocene in South Africa, the Cave of Hearths, Limpopo province. The corollary is that the symmetry present in the Cave of Hearths Large Cutting Tools can be studied with some measure of confidence: suggestions are made regarding the nature of tool typologies and the knappers' ultimate focus on tip shape and utility.

Keywords
ANOVA, by-eye, Cave of Hearths, Flip Test, Large Cutting Tools, subjectivity, symmetry

Introduction
The starting point of this paper is McNabb, Binyon and Hazelwood’s (2004) “The Large Cutting Tools from the South African Acheulean and the Question of Social Traditions”. Despite my disagreement with several aspects of their theoretical stance (Underhill-Stocks 2006; see also Palmer et al. 2005 for a similar view) it is their ‘by-eye’ symmetry methodology which is the focus here. Although newly designed their methodology received little attention in the subsequent comments. Only Machin and Mithen (2004) really focus on the methodology, criticising the symmetry measure from which McNabb et al. ultimately make their suppositions about symbolism. What Machin and Mithen objected to most strongly is the assumed subjectivity involved. However, their dissatisfaction is based on this assumption, with no investigation into the proposed subjective ambiguity within the method. Presented here is a statistical analysis of the by-eye symmetry methodology, focusing on potential levels of subjectivity within it. Ultimately it is clear the method does indeed contain a level of subjectivity which strips it of any analytical value. Following a brief critique of available alternatives the results of an examination utilising Hardaker and Dunn’s (2005) Flip Test is presented to facilitate discussion of the levels of symmetry present within the Cave of Hearths Large Cutting Tools.

A focus on symmetry is important in the present climate of burgeoning cognitive archaeology, due to its accepted relationship with cognitive development (McBeath et
al. 1997; Tversky and Lee 1998); as such it may allow communication between the archaeological record and cognitive sciences, a partnership which should elucidate the development of hominin cognition (Wynn 2002). Through studying changes in both standardisation and symmetry it is believed that one can trace the origin and evolution of the ‘human’ mind (Nowell et al. 2003). One important aspect not considered here is the influence of any re-sharpening and reuse on an artefact’s eventual symmetry. However, McNabb et al. (2004: 668) assert that re-sharpening was “not particularly important to the Cave of Hearths knappers” with only the bare minimum of secondary working present. The local abundance of raw material makes it unlikely that any of the artefacts were curated, and so any re-sharpening will have little effect on the overall patterns revealed.

The By-Eye Symmetry Methodology
The bilateral symmetry measure established by McNabb et al. (2004) (see also Sinclair and McNabb 2005) involves dividing each tool into equal thirds bisected along the long axis, which is established through the presence of a clear convergence, or by passing through half of the maximum width. This creates six sections, three either side of the bisecting line: the tip, medial and basal portions. To establish the extent of any symmetry these sections are imagined folded around the long axis and scored via a binary yes/no system. As each tool consists of three yes/no scores it can be assigned a category based on the eight possible combinations. In addition McNabb et al. (2004: 658) felt it necessary to examine three further relevant classes; near-symmetry was scored for those tools which where not manifestly symmetrical but still retained a “sense of balance”, defined as the knapper evidently paying attention to “the distribution of volume around the midline when seen in planform”. Visually distinctive features, which occur in parallel on opposite edges, were noted as relevant to overall symmetry; for example, notches which initiate the edges sweep toward the tip. Finally, they marked profoundly asymmetrical tips, which evidently resulted from working, as it was the knapper’s decision whether to impose clear bilateral symmetry or not.

Intuitively this appears subjective, which is what Machin and Mithen (2004) object to. Unfortunately, they appear to misunderstand the overall aim of the methodology. Whilst remaining acutely aware of the need to extract patterns from infinite variability McNabb et al. attempted the construction of a simple methodology to enable the re-analysis of the huge number of artefacts present in the South African Acheulean, in order to facilitate the unification of the evidence and allow more effective study to progress. As they themselves state, “we aimed at simple comparisons of assemblages of large cutting tools on an intrasite as well as intersite basis” (McNabb et al. 2004: 656). However, it was patently vital to maintain a level of simplicity to enable the methodology’s uptake. Unfortunately, whilst by-eye judgements allow for a swift turnover in the number of artefacts which can be studied, it does seem unlikely that two unconnected researchers could replicate the same results. Its reliance on pure observation potentially exposes a minefield of biases (Adams and Adams 1991). However there is no proof that this particular method actually does suffer from this. What follows is an investigation partly extracted from the author’s MA dissertation (Underhill-Stocks 2006).
Subjectivity of the By-Eye Measure

To investigate the possibility of observer bias eight independent participants analysed digital images of 50 bifaces taken at random from the Pinel 6 sample in the Marshall et al. (2002) database. A larger sample base would have been desirable but the ad hoc distribution of the test was precluded by the wish to retain a variable demographic: five were graduates (four in archaeology and one in health sciences), three of which were trained in Lithic Analysis, with ages ranging from 24 to 58. The biface images were prefaced by the McNabb et al. symmetry methodology, extracted verbatim from their paper with no further instruction given to any of the participants. Each example was accompanied by a basic table for recording the binary yes/no scoring, in addition to near-symmetry, features promoting symmetry, and tip asymmetry. As an interesting aside even those not trained in lithic analysis, or archaeology, were startled by just how subjective the test felt, yet a belief in subjectivity is no proof of it, which is the entire point of this particular examination.

The results were statistically analysed utilising an Analysis of Variance (Miller 1986; Thayer and Turner 2001). Although SPSS (Version 14.0 used here) has the ability to run this test, it is necessary to understand the process in order to make use of its results. This method examines the results through two groupings: each participant’s 50 results, and each individual artefact’s 8 results generated following the McNabb et al. method. The null hypothesis, that the test is not affected by observer bias, predicts that regard less of participant the artefact’s score should be the same, producing a variance ratio of 1:1 (F distribution). To calculate this, one must first know how the question itself limits the possible variance (degrees of freedom or df). As there are two groups each will affect this in its own way (see Sheskin 2004). Next one must know how each result deviates from the group mean. The calculation \( \sum (X_i - \bar{X})^2 \) for each group will give the population variance. These groups are then divided to gain the F distribution. If the null hypothesis is correct this should equal one, or a confidence limit of .95, representing that the result is statistically significant (alpha or \( \alpha \) value).

Results of Subjectivity Study

The following table represents the full Analysis of Variance for the test on McNabb et al.’s symmetry methodology. This result displays an alpha value of 0.001 clearly revealing that the variation cannot be chance alone: each participant’s subjectivity must have affected the artefact scores, confirming that the bias of the test undermines its ability to reveal true patterns.

It was deemed important to retain a variable demographic to investigate the mitigation of subjectivity. Although not an ordinal scale the mean of each participant’s symmetry total for the assemblage shall be used as a proxy for differences in their perceptions. Foremost in importance is the presence or otherwise of lithic analysis training: a mean symmetry for the assemblage of 3.75 ± 2.23 was established for those with training, compared with 3.95 ± 2.24 for those without. Therefore, as confirmed by an Analysis of Variance there was no statistically significant difference between the two groups. Further examination reveals there was in fact no pattern. The largest disparity exhibited is between participants with a general academic background and those without; a mean
<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TIP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>11.23</td>
<td>7</td>
<td>1.604</td>
<td>7.104</td>
<td>0</td>
</tr>
<tr>
<td>Within Groups</td>
<td>88.52</td>
<td>392</td>
<td>0.226</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>99.74</td>
<td>399</td>
<td>0.226</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MEDIAL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>24.898</td>
<td>7</td>
<td>3.557</td>
<td>19.67</td>
<td>0</td>
</tr>
<tr>
<td>Within Groups</td>
<td>70.9</td>
<td>392</td>
<td>0.181</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>95.798</td>
<td>399</td>
<td>0.181</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BASAL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>15.79</td>
<td>7</td>
<td>2.256</td>
<td>10.65</td>
<td>0</td>
</tr>
<tr>
<td>Within Groups</td>
<td>83</td>
<td>392</td>
<td>0.212</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>98.79</td>
<td>399</td>
<td>0.212</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>303.47</td>
<td>7</td>
<td>43.353</td>
<td>10.02</td>
<td>0</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1696.28</td>
<td>392</td>
<td>4.327</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1999.75</td>
<td>399</td>
<td>4.327</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Results of the Analysis of Variance for the test on McNabb et al.'s by-eye symmetry methodology.
of 4.08 and 3.53 was found respectively (α = .016). All of this points to the inescapable conclusion that there is no way of mitigating the subjectivity of by-eye symmetry judgements.

Alternative Method for Symmetry Study
Having established that the by-eye method is not reliable, the question becomes how to best measure symmetry whilst retaining the necessary simplicity. Few methods exist as until recently symmetry seems to have been assumed (Saragusti et al. 1998, Wynn and Tierson 1990), although much weight is often put upon this assumption. At present there appear to be three viable alternative methods: Wynn and Tierson (1990), Saragusti et al. (1998, 2005) and Hardaker and Dunn (2005). Wynn and Tierson’s method involves measuring 22 polar co-ordinates from the midpoint of the long axis to the edge of an artefact’s drawn outline. It focuses attention on the tip area so as to identify strictly typological distinctions, yet this is the least worked section of the Cave of Hearths assemblage. In addition, it seems redundant to reduce an artefact to 22 points, 20 if scaled to a standard length (McPherron 2000), when a digital image can represent the full outline. More fundamentally this technique is intended as a replacement for the standard measurements of the Roe methodology, yet the latter has already been adopted in African studies (Clark 2001, Leakey and Roe 1994).

More recently, the Continuous Symmetry Measure (CSM) has been proposed by Saragusti et al. This measure is borrowed from chemistry (Zabrodsky and Avnir 1995) and works by establishing “the minimal distances that the vertices of a shape have to undergo, in order for the shape to attain the desired symmetry” (Saragusti et al. 1998: 819). Although this test has primarily been trumpeted by Machin and Mithen (2004) and Machin et al. (2005, 2006, 2007) they have the luxury of being able to feed digital images into software devised for the Marshal et al. (2002) biface database. Unfortunately this software is not readily available, and trying to compute the measure by hand is far beyond most non-statisticians. Without a freely available computerised aid the CSM is far from suitable for simple implementation in order to clarify the already murky South African Acheulean (see also Hardaker and Dunn 2005).

The only alternative available at present is the Flip Test (Hardaker and Dunn 2005) which shares much with the CSM. Both measure deviation from true bilateral symmetry through ‘folding’ one edge over, or in the case of the CSM co-ordinates representing the edge, to establish how far the opposing edge differs. Fortunately the Flip Test was devised to measure the symmetry present in Acheulean bifaces, providing an ‘index of symmetry’. It follows the same principle as the by-eye judgement but utilises digital images; by automatically rotating the image of a tool to reflect its true long axis and then folding the image it is able to superimpose the outlines in order to reflect the deviation from true symmetry (Hardaker and Dunn 2005). The program then supplies a numerical expression of this symmetry using the formula $\frac{500(A)}{(H+W)^2}$ (where $A =$ asymmetrical pixel count (that is the number of pixels between the imposed outlines)) (Fig. 1). In the author’s opinion it is the simplicity of this measure, combined with free ac
cess to the program which makes it the best available method for the South African evidence; although as with any use of digitised images it does necessitate a certain amount of preparation.

The images of the Cave of Hearths artefacts utilised here are available from McNabb et al. (2003). Unfortunately, these pictures were not intended for an examination of this nature, their sole purpose being to illustrate variability (J. McNabb, pers. comm.). Fortunately they are large enough to avoid the problem of distortion with the centre of photographic images being less distorted than the edges (McPherron and Dibble 1999). Aspect ratio is not an issue; any error will be replicated across all the images, as all were taken by the same camera (J. McNabb, pers. comm.) and as the method is comparative only to itself, producing an otherwise arbitrary figure, the exactness of the digital images, whilst problematic, should not distort the essential validity of the test. One unavoidable issue is that the pictures were not taken perpendicular to the artefact, occasionally creating shadows around one edge of the tool, a problem it is impossible to rectify without re-photographing the artefacts. However, every effort was made to strip any shadows back to the visible edge, to mitigate this problem as far as possible (all image manipulation done through GIMP 2.2).

The results the Flip Test generates do require further investigation; it would be useful to employ the test on experimental pieces, with skill levels controlled for. However, Hardaker and Dunn (2005) supply a table to aid in interpreting the Flip Test scores, ranging from one, ‘virtually perfect’ symmetry, to six, ‘very low’ symmetry, as derived

![Figure 1: Hardaker and Dunns Flip Test run on biface 31 from McNabb et al. 2003. It can be seen that the image has been rotated to express the long axis.](image-url)
from their experience, which is primarily with English assemblages (T. Hardaker, pers. comm.). Their interpretation of the scores must at present be taken on trust until further experimentation can be done. As an example of how this index differs from the by-eye judgement: biface 31 from the McNabb et al. (2003) database, is classified as six on their scale bearing no symmetry at all. In contrast when rotated and processed in the Flip Test it produces a result of 2.64 (Fig. 1), classed as ‘very high’ on the Hardaker and Dunn scale, a level of symmetry which they suggest represents “exceptional skill”. Whatever the meaning of this result one thing is clear: a more objective examination of this artefact’s symmetry reveals the inaccuracy of the by-eye judgement.

Results of Symmetry Study

After processing the available assemblage with the Flip Test, Hardaker and Dunn’s (2005) proxy guide suggests the Cave of Hearths LCTs should indeed be classed as low in their symmetry (mean = 5.36). They suggest that an index of this range could result from intractable material. McNabb et al. discount this as a contributing factor, with the majority of the tools constructed on fine-grained homogeneous quartzite which is conducive to realising a knappers’ intent. Initially then, this result would seem to confirm McNabb et al.’s (2003) conclusion that symmetry was not a high priority for the knappers at the Cave of Hearths.

Breaking these results down further reveals some interesting patterns. McNabb et al.’s bifaces have a mean symmetry of 4.55 and their cleavers 5.67, with even the most asymmetrical bifaces being more symmetrical than the most symmetrical cleavers (although see below); symmetry usually increases with working. Indeed artefacts constructed on flake blanks are less symmetrical (mean = 5.60) than those on indeterminate blanks, clasts, or cobbles (mean = 4.68). What this suggests is that the level of symmetry being imposed on blanks was actually rising when more ‘choice’ was included (the isochrestic zone (Sackett 1982, 1986; see also Isaac 1972); it is by no means an a priori side effect of bifacial working or reduction (Noble and Davidson 1996; although see Coventry and Clibbens 2002; ), requiring as it does a greater sense of purpose (Isaac 1976; see also Kohn and Mithen 1999). However, quartzite tools have a lower mean (5.32 ± 2.21) in terms of symmetry than the non-quartzite (5.37 ± 2.35), albeit only slightly, and the proximity of these results could be taken to support an idea that end products were targeted, regardless of raw material. Indeed, as shall become clear it is the tips which appear to be the focus of the Cave of Hearths artefacts, and they are consistently the least worked aspects of the assemblage. The implication is that the cognitive schemas for the construction of both bifaces and cleavers did in fact contain an aspect of symmetry, but that McNabb et al. are correct in concluding that it was not a high priority.

To reveal more from these results requires utilising the tip shapes established and explained by McNabb et al. (2004). The divergent tips are consistently more asymmetrical (mean = 6.16) than convergent tips (mean = 4.65), possibly indicating the lower working index (mean scars = 18.7 versus 23.3 respectively). However, the idea that simple reduction will lead to symmetry is disputed here by the “divergent with square tips” which become less symmetrical as their working index increases. Given that there was no limit of good blanks for their creation (J. McNabb, pers. comm.) the implication
must be that symmetry was even less relevant in these particular types, being passed over by the knapper for other, more pressing concerns. The convergent forms of these square-tipped pieces, which are relatively heavily worked at the tip (mean scars = 12.9), implies that the shape was targeted, supporting an implication of specific functions. The other tip types worthy of note are the “markedly convergent” and “convergent with an oblique tip”: in both, symmetry increases exponentially as working extent increases. The former is to be expected given that these are typologically only bifaces, yet the latter obliquely tipped type is of more interest, being a cleaver form.

The suggestion here is that the oblique-tipped examples be considered as one group regardless of divergence or convergence. They appear to display the imposition of the idea of an oblique tip, no doubt required by the task the tool was created for. When working on the convergent artefacts is low, defined here arbitrarily as below 16 scars, their symmetry index is closer to that found in the divergent forms. When considered together as a singular group of obliquely-tipped artefacts one can see a change in symmetry connected solely with the extent of working. Further evidence for this

![Boxplot showing the correlation of interquartile ranges for the square and obliquely tipped tools.](image)

**Figure 2:** Boxplot showing the correlation of interquartile ranges for the square and obliquely tipped tools.
continuum can be found in the standard deviations of each group’s symmetry index (Fig. 2). This measure illustrates the range of evident variability, revealing that the top of both the convergent forms’ third interquartile ranges match the corresponding divergent forms’ median levels. Consequently it could be argued that the proximity of the divergent square forms’ median to the bottom of the divergent oblique forms’ first interquartile might imply a continuum between these two groups. However, it could equally be caused by error in typological diagnosis which would explain why the convergent oblique’s median is so far from the bottom of the divergent oblique’s first interquartile, even though this would still fit well in the convergent form’s third interquartile range. This is an important suggestion: If it is true that the tip shapes, regardless of whether divergent or convergent, were the same in the knappers mind, then the consequences for cognitive behaviour are radical, particularly when it is remembered that working is largely on the medial and basal sections of the tools rather than the tips.

**Discussion of Symmetry Results**

Typologically speaking cleavers are normally identified if the width of the bit is more than half the width of the tool (Leakey and Roe 1994). The advantage of McNabb et al.’s approach, one of observation-based subdivisions where parallel or divergent tips indicate a cleaver and converging tips indicate a biface, although still including arbitrary subdivisions, is the loss of some of the inflexible boundaries which accompany pigeon-holing variable tools into invariable categories. However, the suggestion that the square and obliquely tipped artefacts at the Cave of Hearths be considered as singular tool types, regardless of their convergent or divergent nature makes this classification unsatisfactory. As White (2006) suggests for the British assemblages, cleavers are one aspect of continuous biface variation. However, rather than deny their independent utility the presence of the idea of a cleaver bit on these pieces suggests they should in fact all be considered as cleavers regardless of the bit width, creating five groups from McNabb et al.’s initial seven: oblique, square, very convex, generalised and markedly convergent. This is further supported by the assertion that the Cave of Hearths shows “a propensity for tips to converge whether by design or otherwise” (McNabb et al. 2003). The problem of pigeonholing is that, by definition, structured criteria are required which do not necessarily describe the variability evident and lead to understandings of ‘idealised types’ which are actually extremely rare (Adams and Adams 1991).

**Conclusion**

The system McNabb et al. devised does allow for the retention of a certain amount of the evident variability, yet it is also not completely satisfactory. This is due not only to the inherent problems of observer bias but also to the suggestion above that there was in fact more of a limit to design in the overall shapes of tools than they identified. Primarily, there appears to have been knowledge of what tip shape was needed to do which job. Indeed, one of the conclusions of McNabb et al. (2004) was that tip shape, probably more so than full outline, was of primary importance to the manufacturers of the Acheulean in South Africa (Deacon and Deacon 1999). The problem of creating a rigorous and relevant method for typing Large Cutting Tools remains, yet as the eventual use is the initial cognitive cue to their creation it seems likely that an answer will be discovered in the analysis of technological characteristics and the
initial aspects of the cognitive chaîne opératoire, rather than any strict typological end product.

It appears that at the Cave of Hearths “the approach to flaking bifaces and cleavers is essentially the same” (McNabb et al. 2003). However, the author’s re-analysis (Underhill-Stocks 2006), concludes that blanks at the site were chosen due to the inherent properties they offered, generally allowing the least work to be done to achieve a functional yet somewhat targeted end product. Quartzite was the primary raw material and non-quartzite appears to have been either less well understood or harder to actually flake, having an, albeit very slightly, lower mean symmetry index than quartzite’s (see above), yet a higher scar count (25.6±9.8 against 18.4±8.4). It can be suggested with some confidence that when the knapper was presented with more options they could and did impose a certain level of symmetry, although the suggestion must remain that the focus at the Cave of Hearths was on making these tools functional.

Importantly the quest to impose a similar symmetry, planform and isometry to pieces at the Cave of Hearths suggests that regardless of the raw material the knappers were seeking a more or less consistent target result. They were employing a cognitive schema, generated by the demands of the task at hand, and it seems clear that the schema being applied to the less tractable raw materials was very similar to the better understood ones; they were apparently not being exploited for their own inherent properties but treated like hard quartzite (J. McNabb pers. comm.). Any symmetry imposed was apparently embedded in the cognitive schema. Rather than being a priority, it was just residual knowledge attached to a particular aspect of the schemata.

To conclude, it can be seen that despite a commendable effort to create a new methodology for the analysis and unification of the South African Acheulean, and although most of their straightforward technological methods seem sound, the by-eye symmetry measure of McNabb et al. is simply not robust enough to utilise. This is because it was not possible to mitigate the problem of observer bias. The present author proposes the Flip Test of Hardaker and Dunn (2005) as a better replacement and although more experimentation is required it is both accurate and easy to implement. The result of applying this method on the Cave of Hearths Large Cutting Tools reveals that the, more robust, index independently confirmed the results of the by-eye symmetry test, despite its problems. However, the more accurate measure of symmetry has enabled the present study to suggest that in the mind of the Acheulean knappers at the Cave of Hearths there was less distinction in tool types than many researchers have suggested, with knappers’ focus primarily being on tip shape combined with ease of manufacture.

Acknowledgements
Specifically Dr John McNabb, for his inspiration and allowing access to his personal data, but primarily for not being too proud to object to this critique. The symmetry test participants: De, Em, Ando, Joan, Lara, Jackie, Karen and Tony. The anonymous referees who commented on the draft of this paper. I must also thank my parents for helping us during our transition to Southampton and my fiancée for accepting the madness that is me with an idea I can’t put down. My research is most graciously part-funded by the Right Honourable Robert Underhill FCA BSc(Hons) and family.
References


