In this pilot study we explore the relationship between bipolar flake shape and stone raw material differences. We conducted a morphometric analysis of 40 quartzite and 40 basalt experimentally replicated specimens from Olduvai Gorge, Tanzania. We carried out two sets of statistical analyses to investigate the shapes of quartzite and basalt bipolar flakes. Our first set of analyses focused on statistical comparisons of five morphometric variables recorded on the two samples of bipolar flakes. For the second set of analyses we used principal components analysis (PCA) to investigate shape differences among the quartzite and basalt bipolar flakes. Our results suggested that there are no significant differences amongst the quartzite and basalt samples. These results are consistent with the hypothesis that the close association between quartzite and bipolar reduction during the Lower Paleolithic at Olduvai Gorge is due to a single motivating factor, namely, expediency. However, we emphasize that more experiments altering test parameters and variables are needed, as are direct quantitative comparisons between experimental and archaeological datasets. Future experiments may uncover additional motivating factors for the prominent quartzite-bipolar reduction found at Olduvai Gorge.

INTRODUCTION

Bipolar reduction, the technique of producing stone flakes via smashing a core or nodule on an anvil with a percussion hammer, can be found in nearly every time period and geographic location in which hominins made stone tools (e.g. Geier 1990; Hayden 1980; Hiscock 1996; Honea 1965; Jeske 1992; Jeske and Lurie 1993; Kuhn 1995; Kuijt and Russell 1993; Mitchell 2000; Mithen et al. 2001; Shott 1989, 1999). Observations of primate nut-cracking have led some researchers to suggest that a better understanding of the bipolar (anvil) technique may shed light on the origins of primate technology and the tools used by the last common ancestor of humans and chimpanzees (Carvalho et al. 2009;
Carvalho and McGrew 2012). Other researchers have investigated links between the increased relative frequency of bipolar reduction, time budgets, and climate change (Eren et al. 2013); between bipolar reduction and mobility (Eren 2010; Goodyear 1993); between bipolar reduction and the production of intended products (Berman et al. 1999); between bipolar reduction and raw material “constraints” (Barham 1987); and between bipolar reduction and tool production efficiency (Diez-Martin et al. 2011; Morgan et al. 2015). While the relative frequency of the bipolar technique through time and space suggests that its adoption and use was context-specific, the number of proposed motivating factors for using bipolar reduction rather than freehand percussion in different times and places is consistent with the idea that the technique was independently invented (evolutionary convergence) several times over the past 2.6 million years of hominin stone tool-making. Recurrent convergent evolution of bipolar reduction may also be likely due to its simplicity: unlike freehand percussion, relatively little to no skill level is required to successfully carry out bipolar reduction.

In the Lower Paleolithic of East Africa bipolar reduction is prominent (Carbonell et al. 2009:28; Hovers 2003; Ludwig and Harris 1998:90; Merrick and Merrick 1976:579; Mgeladze et al. 2011:592–93; Potts 1988:245; Schick and Toth 1994:120–22). Within the specific context of Olduvai Gorge, Tanzania, Gurtov and Eren (2014) recently explored the prominent association of bipolar technology and quartzite (Diez-Martin et al. 2009:281, 2010:377, 2011:690–91), and the lack of association between bipolar technology and basalt, using an experimentally replicated dataset consisting of both of these locally available toolstones. Rather than attribute the close technology-toolstone association between quartzite and bipolar reduction to the usual subjective assessments of stone raw material properties, Gurtov and Eren (2014) linked the quartzite-bipolar connection at Olduvai to the comparative economic benefits offered by quartzite-bipolar reduction relative to non-quartz bipolar reduction (see also Braun et al. 2009). Their pilot results consisted of two principle findings. First, they found that bipolar reduction on quartzite was significantly more “expedient” than it was on basalt. “Expediency” was quantitatively defined in their study as the number of flakes produced per original core mass (g) per unit time (seconds). It is noteworthy that quartzite-bipolar reduction was significantly more expedient than basalt bipolar reduction not only when flake production was considered in relation to core mass and time concurrently, but also when flake production was considered in relation to each of these factors in isolation. Second, Gurtov and Eren (2014) found that there was no significant size difference between flakes produced via quartzite-bipolar reduction versus those produced via basalt bipolar reduction, nor was there any difference in the length of flake cutting edge. In light of these results, they inferred that the bipolar reduction of quartzite at Olduvai could be attributed to increased production expediency rather than the morphometric attributes — flake size and length of cutting edge — of bipolar flake products of either raw material.

In this paper, we build on the results of Gurtov and Eren (2014) by investigating bipolar flake shape. Our analysis below falls under the general study of lithic artifact “morphometric comparative anatomy,” with the explicit goal of furthering “our understanding of the many dynamics structuring Palaeolithic variability and technological change” (see Lycett 2009:88–89). Specifically, we compare the shapes of quartzite and basalt bipolar flakes to determine if raw material differences translate into bipolar flake shape differences, and, if so, whether those shape differences would have potentially influenced raw material selection amongst prehistoric tool-makers. In other words, we are “dissecting” flake shape via quantitative assessment of shape attributes to understand what differences exist amongst bipolar flakes made on distinct raw materials, and evaluating whether any differences could be functionally significant (e.g. Buchanan et al. 2014; Eren and Lycett 2012).

There are four potential outcomes of this analysis. The first is that there are no significant differences between quartzite and basalt bipolar flake shape. The second is that there are significant differences but the differences cannot be attributed to any sort of functional superiority for either raw material. The third is that there are shape differences that suggest that quartzite-bipolar flakes are functionally superior to basalt ones. The fourth is that there are shape differences that suggest that basalt bipolar flakes are functionally superior to quartzite ones. The first two outcomes are consistent with the hypothesis that expediency was the predominant factor in the Lower Paleolithic association between bipolar reduction and quartzite at Olduvai.
Gorge. The third outcome would be consistent with the hypothesis that there were multiple motivating factors for the hominin use of quartzite over basalt in bipolar reduction. Results consistent with the fourth outcome would suggest a conclusion similar to the first in that “expediency” was a driving factor in the quartzite-bipolar connection, but with a further corollary that flake production expediency was so important to Lower Paleolithic hominins at Olduvai that they were willing to forgo “superiorly” shaped basalt bipolar flake products in exchange for bipolar expediency on quartzite.

**MATERIALS AND METHODS**

The experimental dataset used in this study was produced from 30 nodules, 15 quartzite, and 15 basalt, procured locally from Olduvai Gorge, Tanzania. The quartzite was procured from the Naibor Soit source, approximately 3.5 km to the north of the confluence of the two gorges and the basalt was procured from lava beds in close proximity to FLK. Additional specific details of the experimental nodules can be found in Gurtov and Eren (2014:287–88). From the 15 quartzite nodules 368 “usable flakes” (>2.5 cm in maximum dimension) were produced, while 357 usable flakes were produced from the 15 basalt nodules. For the shape analysis presented here 40 flakes were randomly selected from each raw material set (total n = 80) using the online tool Research Randomizer (http://www.randomizer.org/form.htm). Five morphometric variables were recorded in millimeters on each flake: maximum dimension, maximum width, width at 50 per cent of maximum dimension, maximum thickness, and thickness at 50 per cent of maximum dimension. Width and thickness were recorded orthogonally to maximum dimension. These five variables were selected because they could be easily applied in absence of technological attributes like platforms or bulbs of percussion, which may occur twice (hammerstone and anvil ends of flakes) or not at all (chunks and splintered pieces) on the products of bipolar reduction.

We conducted two sets of statistical analyses to investigate the shapes of quartzite and basalt bipolar flakes. Our first set of analyses focused on statistical comparisons of the five morphometric variables recorded on the two samples of bipolar flakes. Prior to the analyses, we determined if the sample distributions of each of the variables conformed to an underlying normal distribution using Kolmogorov–Smirnov tests. All of the variables were normally distributed. We then conducted two sample t-tests on the variables. In addition to the t-tests we report the results of Levene’s test for the equality of variances. Because multiple tests were conducted, we used Benjamini and Yekutieli’s (2001) method of significance-level correction to reduce type-I error rates. We employed this method rather than the better-known Bonferroni correction because it has been shown to balance the reduction of type-I and type-II error rates better than Bonferroni correction (Narum 2006). We conducted these tests to determine if the quartzite and basalt bipolar flakes differ in any of the five variables.

For the second set of analyses we used principal components analysis (PCA) to investigate shape...
differences among the quartzite and basalt bipolar flakes. We used PCA because it is generally accepted to be capable of decomposing form into size and shape by researchers who work with morphometric data (e.g., Strauss 2010). In studies in which PCA is applied to morphometric data, it is usual for the first principal component (PC1) to reflect size variation among the taxa and for the other components, which are orthogonal to and therefore uncorrelated with PC1, to reflect shape variation among the taxa.

After subjecting the data to PCA, we retained the principal components (PCs) that constituted more than 5 per cent of the variation in the dataset. The first three PCs met this criterion and were used in the analyses. Next, we tested if the retained PC scores conformed to an underlying normal distribution using Kolmogorov–Smirnov tests. Because all of the PC scores conformed to a normal distribution, we opted to use t-tests to investigate differences in PC scores among the two samples of bipolar flakes. As in the first set of analyses, we modified the significance level to account for multiple tests using Benjamini and Yekutieli’s (2001) method. In line with the standard interpretation of morphometrics-derived PCs discussed above, we expected that any differences in the shape of basalt and quartzite-bipolar flakes will be expressed in the second and third PCs (Figure 1).

All statistical tests were carried out in SPSS version 20.0.

### RESULTS

Quartzite and basalt bipolar flakes had equal variances for all five variables according to the Levene’s test for the equality of variances (Table 1). Two-sample t-tests revealed that none of the variables were significantly different after applying Benjamini and Yekutieli’s (2001) correction method (Table 1).

The results of the PCA conducted to investigate differences in the shape of quartzite and basalt bipolar flakes are summarized in Figure 2 and Table 2. As indicated above, three PCs met the criterion for retention, these components account for 99.19 per cent of the overall variation in the dataset. As expected, PC1 accounts for the majority of variation in the dataset (81.53 per cent). The large magnitude and consistent direction of the loadings of the five characters supports the assumption that PC1 represents size variation. The other two PCs capture aspects of shape variation, as expected. The loadings on PC2 (10.31 per cent of the variation) indicate a changing relationship between the maximum dimension of the flakes and the other measures (maximum width, width at 50 per cent maximum length, thickness at 50 per cent maximum length, and maximum flake thickness), such that in Figure 2A flakes become shorter, wider, and thicker as one moves up the PC2 axis. The loadings on PC3 (7.34 per cent of the variation) indicate a changing relationship primarily between the width measures (maximum dimension maximum width at 50 per

### Table 1. Results of Levene’s Tests for Equality of Variances and Two Sample t-Tests with 78 Degrees of Freedom Between 40 Basalt and 40 Quartzite-Bipolar Flakes for 10 Morphometric Variables. Significance Assessed Using Benjamini and Yekutieli’s (2001) Method (Alpha Level for Five Tests is 0.02190)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Levene’s test for equality of variances</th>
<th>t-Test for equality of means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F  Sig.</td>
<td>t</td>
</tr>
<tr>
<td>Max dimension (mm)</td>
<td>0.008 0.927</td>
<td>-0.446</td>
</tr>
<tr>
<td>Maximum flake width</td>
<td>0.100 0.753</td>
<td>-0.879</td>
</tr>
<tr>
<td>Width at 50% max length</td>
<td>0.195 0.660</td>
<td>-0.681</td>
</tr>
<tr>
<td>Thickness at 50% of max length</td>
<td>1.457 0.231</td>
<td>-2.293</td>
</tr>
<tr>
<td>Maximum flake thickness</td>
<td>0.016 0.900</td>
<td>-1.774</td>
</tr>
</tbody>
</table>
cent maximum length) and the thickness measures (thickness at 50 per cent maximum length and maximum flake thickness), such that in Figure 2B flakes become thicker and narrower as one moves up the PC3 axis.

Figure 2A shows considerable overlap among the quartzite and basalt bipolar flakes along the PC1 axis. Not surprisingly, a $t$-test of the scores from PC1 indicates that there is no significant difference in the size of quartzite and basalt bipolar flakes ($t = -1.00, df = 78, P = 0.32$). This is consistent with the findings of Gurtov and Eren (2014). Similar to PC1, the quartzite and basalt flakes show little separation along the PC2 axis (Figure 2A). A test of PC2 scores between quartzite and basalt flakes confirms that there is no significant difference ($t = -0.93, df = 78, P = 0.36$). Figure 2B also shows considerable overlap

![DISSECTING QUARTZITE AND BASALT BIPOLAR FLAKE SHAPE](Lithic Technology 2015, Vol. 0 No. 0, 1–10)
among the quartzite and basalt flakes along the PC3 axis. PC3 scores also were not significantly different between quartzite and basalt flakes ($t = -1.93$, df = 78, $P = 0.058$).

**DISCUSSION**

The experimental analysis and results presented above were concerned with whether the close association between quartzite and the bipolar technique at Olduvai Gorge could be potentially accounted for by a single motivating factor (i.e. flake production expediency, as suggested by Gurtov and Eren 2014), or multiple motivating factors (i.e. flake production expediency and flake shape). Toward this end, we conducted a comparative quantitative assessment of the shape of experimentally produced quartzite and basalt bipolar flakes. Our analysis did not reveal any significant differences in quartzite and basalt bipolar flake shape, which, based on the limited experimental analyses conducted thus far, is consistent with the hypothesis that the close association between quartzite and bipolar reduction during the Lower Paleolithic at Olduvai Gorge is due to a single motivating factor.

In some respects, the experimental results presented here are surprising. Gurtov and Eren (2014) noted that the basalt bipolar experimental reduction was extremely difficult. They proposed that it was the toughness of this raw material that prevented, on average, 31 per cent of core mass from being exploited because a crack could not be initiated (Gurtov and Eren 2014:290). Alternately, the local quartzite was extremely brittle and easily fractured. If bipolar reduction on basalt is more likely to occur on relatively thinner nodule margins and edges where a crack is more easily initiated, while knapping bipolar reduction on quartzite can initiate cracks closer to a nodule’s center, it follows that the resulting bipolar flakes from these two raw materials might possess different shapes. Similarly, it might be thought that basalt’s toughness could prevent any successfully initiated cracks from propagating as far as they would on quartzite, again resulting in shape differences between bipolar flakes made from the two toolstones. Neither prediction was supported by our experimental work. Instead, when shape variables are considered separately from size variables, our results suggest that bipolar reduction produces similar flake shapes regardless of raw material. In other words, smashing rocks results in smashed rocks no matter which rocks are smashed. Our results are thus more broadly consistent with a growing set of archaeological and experimental data that suggest internal and external differences of stone do not automatically “determine” stone tool morphology (e.g. Archer and Braun 2010; Buchanan et al. 2013; Clarkson 2010; Costa 2010; Eren et al. 2011, 2014; Sharon 2008; Smallwood 2012).

It must be emphasized, however, that we are not saying more than one motivating factor did not potentially govern the close association between quartzite and bipolar reduction at Olduvai, only that experimental results conducted thus far support a single motivating factor (flake production expediency). It is therefore noteworthy that several researchers have suggested that Olduvai’s Naibor Soit quartzite yields more durable edges relative to other raw materials (Blumenschine et al. 2008; Jones 1994; Tactikos 2005). While to our knowledge this assertion has yet to be experimentally tested, it is certainly plausible. This assertion is also extremely intriguing, given that Braun et al. (2009) have demonstrated the Oldowan hominins at Kanjera South (Kenya) had the capacity to select their raw materials based on durability. A series of flake cutting efficiency experiments along the lines of the recent work of Key and Lycett (2014) and

<table>
<thead>
<tr>
<th>Variable</th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Variation</td>
<td>81.53</td>
<td>10.31</td>
<td>7.34</td>
</tr>
<tr>
<td>Max dimension</td>
<td>0.5617</td>
<td>-0.8273</td>
<td>-0.0110</td>
</tr>
<tr>
<td>Max width</td>
<td>0.5093</td>
<td>0.3514</td>
<td>-0.3712</td>
</tr>
<tr>
<td>Width at 50% max length</td>
<td>0.4958</td>
<td>0.3403</td>
<td>-0.3429</td>
</tr>
<tr>
<td>Thickness at 50% of max length</td>
<td>0.2812</td>
<td>0.1773</td>
<td>0.5936</td>
</tr>
<tr>
<td>Maximum flake thickness</td>
<td>0.3168</td>
<td>0.2120</td>
<td>0.6262</td>
</tr>
</tbody>
</table>
Braun et al. (2009) could potentially tease out the relationship, if any, between edge durability and raw material selection at Olduvai.

While experimental stone tool replication provides “invaluable insights” (Magnani et al. 2014) into prehistoric stone tool manufacturing behaviors as well as the potential motivating factors for those behaviors, no experiment should be “the last word.” Not only should the pilot replication experiments and analyses presented here be repeated, but future assessments should attempt to improve on them by integrating other sorts of analyses (e.g. studies of the archaeological record, the use of machine-flaking experiments, see Lycett and Eren 2013). We see five avenues of future investigation with respect to the particular analyses presented here. First, comparing our experimental dataset with the archaeological record of Olduvai to determine how similar our findings are, and how similar the inferences drawn here remain valid ones. If significantly different, then different inferences must be formulated. Second, this study, as well as Gurtov and Eren (2014), should be repeated with larger samples sizes, while additional tests should alter the protocols and variables. For example, the flake size cut-off in this study and in Gurtov and Eren (2014) was 2.5 cm. Including flakes under this size threshold in future experimental analyses may provide interesting comparisons, especially given that there are ethnographic and archaeological instances of foragers making and using flakes smaller than 2.5 cm (Dibble and McPherron 2006; Hiscock 2015; Shott and Sillitoe 2005). Also, the use of 50° as a threshold for cutting edge angle in both studies might be changed in future tests in light of Key and Lycett’s (2015) recent results. Given that replication of archaeological experiments by multiple researchers is rarely practiced, the mere act of replicating our experiments in order to test our findings will be an important step forward toward a quantitative and scientific archaeology. Third, there is a clear need for better characterization of the raw material properties of Olduvai’s quartzite and basalt toolstone to quantitatively understand how they are different (e.g. Braun et al. 2009; Eren et al. 2014). Fourth, machine-flaking experiments (e.g. Rezek et al. 2011) that investigate the relationship between raw material differences and flake properties (size, shape) may shed additional light on the morphometric differences between quartzite and basalt bipolar flakes, in any. Finally, functional (use) experiments that examine the relationship between functional efficiency and raw material differences will play an important role in supporting or falsifying the potential linkage between bipolar flakes, durability, and raw material selection hypothesized by others (Blumenschine et al. 2008; Jones 1994; Tactikos 2005).

If all of these avenues of future investigation seem like overkill, especially for studying one narrowly focused topic of the East African Lower Paleolithic, that is perhaps because for too long lithic analysts striving toward an archaeological science of lithic technology have been advocates for their one preferred type of analysis (e.g. machine-flaking versus human flintknapping; linear measurements versus “appropriate technological” units of analysis; etc.) rather than unifying around a core set of scientific principles (see Lycett 2011; Lycett and Chauhan 2010; Lycett and von Cramon-Taubadel 2014) that integrate and emphasize the links between these different sorts of analyses, that exploit each analysis’ strengths, and that recognize each analysis’ weaknesses, for approaching particular hypotheses and questions (Lycett and Eren 2013; Mesoudi 2011).

ACKNOWLEDGMENTS

Financial support for this research was provided by the British Academy Grant “Can Olduvai Gorge Shed Light on the Middle Stone Age?” (SG112454) and a Leverhulme Trust Early Career Fellowship (ECF-2011-567). MIE is currently supported by a University of Missouri College of Arts and Sciences Post-Doctoral Fellowship. We are grateful to Justin Pargeter and Hilary Duke for inviting us to participate in the SAA Symposium from which this paper originated. We are also appreciative to the anonymous reviewers and to Grant McCall.

NOTES

1 In a recent article Hiscock (2015) expressed concern that “weak” definitions of bipolar reduction used by us and others (Gurtov and Eren 2014:285; also Eren 2010; Eren et al. 2011; Diez-Martín et al. 2011:692) — which demand minimal procedural similarity for inclusion — have been used over “stronger” definitions. We disagree with Hiscock’s attributions and suggest that the “weak” versus “strong” dichotomy is more precisely characterized as a “broader” versus “more specific” dichotomy. We also suggest that the choice to use a broader or more specific definition depends on the question being asked, and does not preclude the use of the alternative (i.e. broader or more specific) in future analyses. Indeed, the juxtaposition...
of results coming from analyses that use broader versus more specific definitions may itself be interesting or informative depending on the question. In the cases that Hiscock (2015) highlights, the topics of study do not necessarily require a specific definition of “bipolar,” as they are concerned with comparing “bipolar” more generally (knapping with use of an anvil) with “freehand” percussion (knapping without use of an anvil). The choice of a broader or more specific definition also depends on whether researchers have the archaeological resolution to distinguish between different types of bipolar (striking a core directly at 90° versus simply resting a core on an anvil before hitting it). Indeed, it would be irresponsible to use a specific definition of bipolar in archaeological analyses when, to our knowledge, the ability of archaeologists to morphometrically identify and distinguish between different types of bipolar procedures suggested by Hiscock does not currently exist. In terms of the experimental dataset used in this paper and in Gurtov and Eren (2014), bipolar reduction did indeed operationally aim to fit Hiscock’s (2015) specific definition, namely “a hammer strikes a rock which is resting on an anvil, and strikes into it at close to 90° and in line with the point at which the rock is in contact with the anvil.” However, the experimental dataset was created in the field, and we did not possess a video recording or some other means with which we could control for, or confirm later, that indeed every blow was close to 90°, despite that being the knapper’s goal. Further, Hiscock’s definition is vague (i.e., how close to 90° is “close”?) and therefore difficult to operationalize. Given these issues, we felt it was more conservative to use a broader definition of bipolar in this study and in Gurtov and Eren (2014:285), “Bipolar reduction can be defined as a percussion technique in which a stone core is placed on an anvil and struck with a hammer to produce flakes”.

2 It is noteworthy that because the experimental dataset used here was created for another analysis, in some respects its use here acts as a “blind” test. Blind testing has yet to be regularly practiced in stone tool replication experiments. However, it is important to understand that blind testing is merely another type of experimental variable control whose use depends entirely on the question being asked. Sometimes the use of blind and non-blind replication experiments in conjunction may be the most productive experimental strategy. Other times use of “blind” stone tool replicated assemblages (e.g. see discussion of Tactikos (2003) in Eren et al. (2008)) or blind tests may confound results, in some cases it may be the most robust research strategy, while in still other situations it may not matter one way or another.

3 During the revision of this manuscript, an experimental analysis of novice knappers using chert was published by Putt (2014). Her results were consistent with Gurtov and Eren’s (2014) results that bipolar reduction is more expedient than other forms of knapping.

REFERENCES


---

**Notes on Contributors**

Alia N. Gurtov is a Ph.D. candidate at the University of Wisconsin, Madison, WI.

Correspondence to: Alia N. Gurtov, Briggs Buchanan, and Metin I. Eren, Department of Anthropology, University of Wisconsin, Madison, WI 53706, USA. E-mail: gurtov@wisc.edu; briggs@utulsa.edu; erenm@missouri.edu.

Briggs Buchanan is an Assistant Professor at the University of Tulsa, Tulsa, OK.

Metin I. Eren is the University of Missouri College of Arts and Sciences Post-Doctoral Fellow, Columbia, MO.