EXPLORING RETOUCH ON BIFACES:
UNPACKING PRODUCTION,
RESHARPENING, AND HAMMER TYPE

Abstract
Measuring retouch amounts on stone tools has been helpful for understanding human organizational strategies. Multiple retouch indices geared toward assessing retouch amounts on flake tools and unifaces have been developed, but few have been developed to evaluate retouch exclusively for bifaces. For this study, a retouch index was developed and evaluated on an experimental assemblage of bifaces. It is shown that reduction activities on bifaces may create extensive amounts of retouch that are contingent upon a number of factors from both the production and resharpening events that must be taken into consideration before understanding a biface’s life history.

INTRODUCTION
Tool curation has been defined as the relationship between a tool’s potential utility and its actual usage (Andrefsky 2005; Bamforth 1986; Shott 1996), or its “life history” (Eren et al. 2005). This curation concept has been linked to studies of hunter-gatherer organizational strategies in understanding issues of land use, economy, and mobility. For stone tools, retouch amount has been used as an effective measure to assess the degree to which a tool has been curated (for discussion of curation see Andrefsky 2006; Barton 1988;Binford 1973, 1979; Blades 2003; Clarkson 2002; Davis and Shea 1998; Dibble 1997; Nelson 1991; Shott 1989, 1996).

However, assessing retouch amount may not be as universal as we might initially believe. We define retouch as the deliberate modification of a stone tool edge created by either percussion or pressure-flaking techniques (Andrefsky 2005). As such, retouch takes place in the beginning production stages of a tool as well as in the subsequent episodes of resharpening and reshaping of a tool’s cutting edge. Therefore, we would expect the amount of retouch to progressively increase throughout the production and the life of a tool. To evaluate retouch in terms of degree of curation, analysts have created indices that quantify retouch for comparisons of stone tools.

Previous retouch measures have been effective for different kinds of stone tool forms. Barton (1988) and Clarkson (2002) measure retouch on flake tools based upon progressive use of the original flake blank. Kuhn (1990) measured retouch on scraper edges. Andrefsky (2006) and Hoffman (1985) measured retouch on hafted bifaces (see also Eren and Prendergast; Hiscock and Clarkson; Quinn et al., this volume). We feel that North American bifaces represent a different tool type than some bifaces from other parts of the world. Bifaces are stone tools that have two surfaces (or faces) that meet to form an edge around the entire perimeter and usually have flake scars that extend from the edge to the midline of the surface (Andrefsky 2005). North American bifaces tend to undergo a production phase and a subsequent use-life phase (Callahan 1979; Whitaker 1994). In some areas of the world, bifaces are produced from flake blanks as a result of their being used and resharpened extensively (cf. Clarkson 2002). However, we feel that some bifaces do not become bifaces as a result of this use and resharpening process. We feel that some bifaces are shaped by extensive retouch before they are even used. Some bifaces, particularly those in parts of North America, are extensively retouched during their production phase, and thus, the retouch amount on the biface has little or no meaning with regard to curation. We suggest that retouch indices should be specifically tailored to different kinds of tools and that we need to consider the differences between bifacial production and bifacial resharpening after use.

To gather information on biface retouch, we conducted a series of production and use experiments attempting to replicate bifaces similar to those recovered from Chalk Basin, a chert quarry workshop area on the Owyhee River in southeastern Oregon. Our experiment
systematically gathered attribute information from both the bifaces being produced and used and thedebitage resulting from the experiment.

THE EXPERIMENT

Our experimental study involved the production of three “quarry bifaces” made from high-chipping-quality chert. Information on each biface was recorded after six arbitrary production and use-life events. The first two events were arbitrary production events and the last four were associated with resharpening episodes after tool use events. When the biface was reduced by approximately half of its starting weight during the production process, we arbitrarily stopped and collected debitage shatter for that production event. Resharpening episodes occurred when the edges of the biface were retouched enough so that it could be used as a tool with a cutting edge around the entire perimeter. The biface edges were then dulled and resharpened again to create a series of resharpening episodes of each biface.

One of the authors performed all of the flintknapping over a drop cloth using either a hard hammer or a soft hammer percussor, while the other author recovered and numbered each flake as it was removed. All production and resharpening were done with percussion flaking (no pressure flaking). The greatest number of flakes collected was from the first production event, which yielded an average of twenty-nine flakes per biface. This makes intuitive sense, because the biface was reduced by the greatest amount during this episode. The smallest number of flakes collected was from the first resharpening event, with an average of 12 flakes collected for each biface. The resharpening events had the greatest amount of variability amongst all three bifaces. The average number of flakes collected for each biface, after the first resharpening event, was nineteen flakes per event. The cores chosen for the experiment were all roughly the same shape and size (approximately weighing 1,000 g each). However, one biface had about twenty more flakes removed from it during the experiment than the other two bifaces. This was due to the presence of material flaws that had to be removed in order to maintain an effective cutting tool.

After each event, all of the shatter was collected and the biface was photographed and measured. For consistency, all of the measurements were done by just one of the authors throughout the experiment. Initially, all three bifaces were reduced using a quartzite hard hammer to remove most of the cortex from the objective piece. After the first half-life, a silex hard hammer and a soft hammer (i.e., antler billet) were used to shape and thin the bifaces. Gradually, throughout the experiment, the percentage of hard hammer flakes decreased whereas the percentage of flakes made by a soft hammer increased.

DEBITAGE PATTERNS WITH BIFACE PRODUCTION AND RESHARPENING

In a previous study (Wilson and Andrefsky 2006), we explored the variability found in debitage characteristics between biface production and biface resharpening events from the experiment. From 256 proximal flakes analyzed, we found that debitage characteristics were significantly associated with differences in production and resharpening events. Metric variables sensitive to these different retouch activities include maximum length, width, thickness, weight, and platform area (maximum platform width multiplied by maximum platform thickness) (Table 4.1). Nominal attributes that were sensitive to retouch activities were platform type and presence of cortex.

Using platform types previously defined (Andrefsky 2005), we found that flakes made from biface production exhibit more flat or cortical platforms than flakes made during resharpening events (Figure 4.1). Most of the flakes also had dorsal cortex and have a smaller width-to-thickness ratio, heavier weight, and larger platform area (Table 4.1).

In contrast, flakes that are the by-products of resharpening events tend to have more complex and abraded platforms. The flakes produced from resharpening events also weigh relatively less, with a smaller platform area, and have a higher width-to-thickness ratio. Even though the two comparative groups in the debitage study did have some overlap, the average sizes of the two groups were significantly different.

Given the results of differences noted in the debitage attribute analysis, we could expect to see a positive correlation between flake weight and platform area, and also between flake weight and width-to-thickness ratio. Intuitively, if the weight increases, there should
Table 4.1. Comparison of attributes recorded from proximal flakes

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Max</th>
<th>Min</th>
<th>Mean</th>
<th>Std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Production</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (g)</td>
<td>99.5</td>
<td>0.7</td>
<td>12.784</td>
<td>15.935</td>
</tr>
<tr>
<td>Width (mm)</td>
<td>93.8</td>
<td>1.5</td>
<td>41.245</td>
<td>16.615</td>
</tr>
<tr>
<td>Length (mm)</td>
<td>113.1</td>
<td>4.6</td>
<td>42.112</td>
<td>17.916</td>
</tr>
<tr>
<td>Thickness (mm)</td>
<td>25.9</td>
<td>2.8</td>
<td>7.870</td>
<td>4.222</td>
</tr>
<tr>
<td>Platform width (mm)</td>
<td>58.8</td>
<td>7.4</td>
<td>10.499</td>
<td>10.622</td>
</tr>
<tr>
<td>Platform thickness (mm)</td>
<td>16.1</td>
<td>1.0</td>
<td>5.852</td>
<td>3.085</td>
</tr>
<tr>
<td>Platform area (mm$^2$)</td>
<td>946.7</td>
<td>7.4</td>
<td>131.822</td>
<td>138.801</td>
</tr>
<tr>
<td>Width to thickness (mm)</td>
<td>15.37</td>
<td>0.48</td>
<td>5.8859</td>
<td>2.5206</td>
</tr>
<tr>
<td><strong>Resharpening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (g)</td>
<td>5.3</td>
<td>5.3</td>
<td>1.193</td>
<td>1.054</td>
</tr>
<tr>
<td>Width (mm)</td>
<td>44.4</td>
<td>7.2</td>
<td>18.736</td>
<td>7.054</td>
</tr>
<tr>
<td>Length (mm)</td>
<td>65.2</td>
<td>8.5</td>
<td>24.315</td>
<td>11.237</td>
</tr>
<tr>
<td>Thickness (mm)</td>
<td>7.5</td>
<td>0.7</td>
<td>2.308</td>
<td>.873</td>
</tr>
<tr>
<td>Platform width (mm)</td>
<td>20.5</td>
<td>2.1</td>
<td>5.363</td>
<td>3.729</td>
</tr>
<tr>
<td>Platform thickness (mm)</td>
<td>5.3</td>
<td>0.4</td>
<td>1.892</td>
<td>.852</td>
</tr>
<tr>
<td>Platform area (mm$^2$)</td>
<td>91.2</td>
<td>1.7</td>
<td>17.693</td>
<td>15.556</td>
</tr>
<tr>
<td>Width to thickness (mm)</td>
<td>22.50</td>
<td>3.85</td>
<td>8.6019</td>
<td>3.0650</td>
</tr>
</tbody>
</table>

FIGURE 4.1. Platform types identified on proximal flakes.

FIGURE 4.2. Platform area of proximal flakes plotted against their weight.

also be an increase in platform area, given that bigger flakes usually have larger platforms, and there should also be a decrease in the width-to-thickness ratio, assuming that the more a flake weighs the larger it should be in size, which is expressed as a ratio. Figure 4.2 displays a scattergram that shows a strong and significant ($R^2 = .4671$, $F = 94.979, p < .001$) relationship between increasing flake weight and platform area with production flakes. From the resharpening episodes, flakes clustered together around the lower weights and smaller platform areas. This relationship was not as strong (Pearson's $r = 0.336$) as with production flakes but was still statistically significant ($R^2 = .1151$, $F = 18.027, p < .0005$).

When flake weights were plotted against the width-to-thickness ratio, it appeared that the weight of the flake increased as the ratio began to decrease (Figure 4.3). On closer examination, however, this was a significant ($R^2 = .0403, F = 6.781, p = .010$) but weak correlation for resharpening flakes. For the production events, there was an insignificant relationship between the variables, where only 0.7% of the variance could be explained ($R^2 = .0074, F = 0.880, p = .350$).
These correlations have shed light on how the weights of production and resharpening flakes relate to platform area and width-to-thickness ratio. There is more variation between the groups in regards to size (width to thickness) and weight, in comparison to the stronger correlation with weight and platform area (i.e., as the weight of the flake increases, so does the platform area for production and resharpening events).

GENERAL BIFACE PATTERNS OF PRODUCTION AND RESharpening

Based upon results gathered from our debitage pattern study, we expected that biface size, shape, and flake removal patterns would also reveal differences between retouch associated with production and retouch associated with resharpening. When graphed, it is apparent that all three bifaces show a continual decrease in both surface area and weight throughout the use-life events (hereafter called use-life events) (Figures 4.4 and 4.5). This is what would be expected given the fact that the use-life events follow a reductive process, resulting in progressively smaller bifaces. However, these data also suggest that the biface use-life events 1 and 2 are responsible for the greatest amount of size reduction and that biface size reduction is significantly less during the resharpening events (3–6).

This pattern is clear when we graph the amount of surface area lost during each use-life event. Even though the amount of total surface area of bifaces progressively decreased during the use-life events, the amount of surface area lost stabilized after the production events 1 and 2 (Figure 4.6). Essentially, the resharpening events (3–6) show very

**FIGURE 4.3.** Width-to-thickness ratio of proximal flakes plotted against their weight.

**FIGURE 4.4.** Total surface area of the bifaces throughout the experiment.

**FIGURE 4.5.** Weight of each biface after each event throughout the experiment.
little lost surface area; the average surface area lost for each biface is roughly 50 cm², compared to about 200 cm² lost during production. This pattern also suggests that there may be some observable differences in biface characteristics between production and resharpening events. However, lost surface area is only effective for discriminating such events in a controlled experimental setting. It is not possible to effectively use such a measure on excavated assemblages, because surface area lost can only be calculated based upon knowing the original size of the biface. However, like the change in debitage attributes, it does suggest that other biface characteristics might help assess differences between production and resharpening events.

RETOUCH INTENSITY

Other studies have shown that retouch intensity has been an effective measure of curation on stone tools (Clarkson 2002; Eren and Pendergast, this volume, Quinn et al., this volume; Hiscock and Clarkson 2005; Kuhn 1990). We suggest that bifaces have a unique production life and use life and thus, retouch amount has to account for these two phases of a biface life cycle. To assess our assumption, we applied Clarkson’s (2002) index of invasiveness to our experimentally produced bifaces.

Each side of the biface was partitioned into eight equal segments (cf. Clarkson 2002), each one accounting for 12.5% of the total area (Figure 4.7). After a reduction event, each segment was given a score of either 0, 0.5, or 1. Segments exhibiting no retouch would receive a score of 0. If the flake patterning was evident but did not reach the midpoint area of the artifact, defined as the arbitrary line from the midline of the biface to the edge, that square would have a value of 0.5. A score of 1 was given to squares where retouch extended from the edge of the biface and past the midpoint area. The scores from each square were added up and then divided by 16 for the average retouch amount, which was the index of invasiveness score. If the invasiveness score was close to 0, the biface would be considered to exhibit little to no retouch. When the invasiveness score approached 1, the tool is classified as being completely retouched.

We found that retouch amount using this technique is not sensitive to resharpening after the production phase. Using this index, the bifaces were scored as heavily retouched after the second use-life event.
FIGURE 4.8. Results of applying Clarkson’s index of invasiveness (2002) to our experimental bifaces.

One of the bifaces even reached a value of one (maximum retouch amount) after the first use-life event. This is interesting because we know the bifaces were never used. However, the index reveals a maximum level of retouch and subsequently a maximum level of curation. This suggests that the index of invasiveness may not be a good indicator of bifacial retouch as it relates to curation, and also that bifaces are produced, used, and resharpened differently than artifacts such as flake tools (cf. Andrefsky 2006). The outcome of this method is not surprising, as Clarkson (2002: 72) does warn about the potential shortcomings of the index of invasiveness when applied to artifacts that have been “fully retouched.” Clarkson’s index was intended for application to bifacially retouched flakes.

RIDGE COUNT RETOUCH INDEX

Clarkson’s index of invasiveness does not adequately segregate biface production from biface resharpening after use. These two use-life events are important in measuring retouch on bifaces. One of the things that intuitively appear to be occurring on the surface of our experimental bifaces is a progressive increase in the number of flake removal scars from early production events to final resharpening events. To explore flake removal scar counts, we developed a retouch index based upon a sample of the flake removal patterns found on the surface of each biface. The average number of ridge counts was used as a proxy for flake removals to derive this index.

To test this retouch index, we collected biface data after each biface use-life event (weight, maximum length, width, thickness, and flake ridge count). The flake ridges were recorded in a systematic way that involved scanning the biface at a high resolution (600 dpi) and then sampling the bifacial surface image using Deneba’s Canvas 8 drafting program. The analysis of each biface image was partitioned using Chris Clarkson’s grid for evaluating retouch invasiveness, which partitioned each side of the biface into eight segments. Once the grid was digitally superimposed on the biface, six 1 × 1 cm squares were drawn on the biface and positioned in the same location after each use-life event (Figure 4.9). Three 1 × 1 cm squares were sampled on each face of
Biface 2 After Production

FIGURE 4.10. Image of one of the analysis squares from one of the experimental bifaces, showing how flake ridges were counted.

light and contrast. The ridges identified on the scanned image were checked on the actual biface to ensure that the lines observed were not biface fissures or ripple marks but actual flake ridges. Once the number of ridges for each square was confirmed, all six ridge counts were added up and divided by six. This resulted in an average ridge count for each biface after each use-life event.

This retouch index was applied to our assemblage of replicated bifaces, with expectations that there would be significant differences between production and retouch use-life events, as seen in the debitage data, and in the amount of surface area lost on the experimental bifaces. The average ridge count associated with each experimentally produced biface use-life event illustrates that the ridge counts increase throughout all of the use-life events before dropping at use-life event 5 during resharpening (Figure 4.11). This pattern reveals some interesting aspects of biface production and resharpening after use. First, the ridge count measure seems to work as an effective tool to assess use-life events from the beginning of the production cycle through the fourth use-life event, and retouch seems to increase as each use-life event increases. However, this progressive pattern ends at use-life event 5, where there is a drop in the retouch index. We
also see that the retouch progression is not markedly different between biface production and biface resharpening, as noted in the debitage data.

Since this was not what we had expected, we began exploring our experimental data to determine what might account for the ridge count drop at use-life event 5. One immediate pattern discovered was that the type of hammer used during the replication experiments gradually changed from hard hammer percussion to soft hammer percussion as the bifaces were progressively retouched. Other studies also suggest that hammer type and density can be important for flake removal (Andrefsky 2007; Cotterell and Kamminga 1987; Dibble 1995; Hayden and Hutchings 1989). Figure 4.12 charts our experimentally derived use-life events against the relative proportion of hard and soft hammer percussion used to remove flakes. The first three events are primarily hard hammer percussion; this changes to approximately 12% during event 4 and down to 2% during event 5, and then it goes back up to close to 30% during event 6. The steep drop in hard hammer percussion from events 3 through 5 and the subsequent rise at event 6 mirrors the ridge count pattern, and suggests to us that the ridge count index is sensitive to the type of hammer used in biface production and resharpening technology.

To explore this relationship further, we plotted the ridge count index and hammer type along with the use-life events (Figure 4.13). The ridge count index for use life events 4–6 is indeed similar to the relative percentages of hard hammer percussion. However, it is also apparent that the ridge count index is sensitive to previous flake removals on the biface. For instance, use-life events 1–3 have high values for hard hammer percussion, yet the ridge count index shows a steady increase from less than 1.0 to over 3.3. Essentially, the ridge count index is increasing as the original nodule is being progressively worked, even though there is minimal change in the percussion technology.

However, we also feel that the ridge count index is related to the existing flake removal pattern on the biface and not solely associated with the type of percussion technology used. For example, Biface 2 in event 5 and Biface 3 in event 3 both have steep drops in the average ridge count (see Figure 4.11). During these particular times of the experiment, these bifaces had irregular flaws in the material that
had to be removed in order to continue to use the biface for usage and resharpening episodes. In doing so, a large portion of the biface surface was removed, including the previous flake ridges, which may have greatly affected the number of flake ridges for particular analysis grids.

SUMMARY AND DISCUSSION

Although incomplete at this point, our analysis shows some interesting trends and potential avenues for further exploration with regard to retouch on bifaces. First, it appears that retouch on bifaces may not be the same as retouch on flake tools. Bifaces are retouched throughout the reduction sequence, even during the production phase. The biface core has to be reduced in a fashion where the edge is continually being modified or retouched. Thus, traditional measures linking retouch amount to curation amount may not be effective for bifaces, because they may have a high retouch score without ever having been used, as illustrated with the application of Clarkson's (2002) index of invasiveness.

Second, overall flake removal amount may be a good indicator of the use life events for bifaces. For instance, our experiment showed that flake removal patterns of biface surfaces tended to increase as the biface was progressively used and resharpened. However, flake removal amount is also sensitive to changes in hammer type. As hammer types change, so does the relative proportion of flake shapes and sizes, which influences the flake removal pattern found on the biface, i.e., raw material flaws or “problem areas.” The hammer type used, soft hammer versus hard hammer, is an idiosyncratic choice that is not a constant. Depending on the skill and technique of the flintknapper, different types of hammers will be used to address or reduce the objective piece into the desired form. The goal of the various flintknappers may be the same but the technique/method will vary from person to person and possibly from stone tool to stone tool (even when the same type of stone tool is being made). This means that flake removal patterns on bifacial surfaces may be effective for interpreting reduction only if hammer type is held constant or can be accounted for in some other way.

Finally, it is also apparent from our data that flake removal amount is not sensitive to changes in biface production events vs. biface resharpening events. Even though these events are clearly visible with debitage characteristics, they are not evident from the surface of bifaces, because the flake removal pattern of bifacial surfaces is produced by a series of multiple technological factors. As previously discussed, these can include flintknapper skill and technique, raw material quality, hammer type, and reduction strategy.

In summary, retouch indices created for flake tools may not be suitable for understanding curation strategies for bifaces. As noted in several other papers in this volume, retouch is particular to different tool types (Andrefsky; Eren and Prendergast; Quinn et al.) and to different tool functions (MacDonald). Retouch does not always equate to tool curation. Retouch is a technique used to shape a tool within the context of tool production, use, and resharpening. All of these contexts must be considered in attempting to quantify tool curation.

REFERENCES CITED


