

2

Measuring Artefact Reduction - An Examination of Kuhn's Geometric Index of Reduction

Peter Hiscock and Chris Clarkson

Abstract

A growing number of techniques have been proposed in recent years to quantify how much retouch has been applied to flakes. This paper reviews the most prominent of these, and evaluates one in particular – Kuhn's (1990) Geometric Index of Unifacial Reduction. This involves a simple experiment designed to explore the performance of the index over a sequence of retouching events for a population of thirty flakes. The results indicate that the index performs admirably in relation to absolute measures of reduction under experimental conditions, and does so especially well in comparison to a number of common alternative techniques.

Introduction

Modern examinations of Palaeolithic artefact assemblages typically depict the complexity of reduction processes and aim to quantify the extent, nature and variability of reduction. Additionally, there is a call to explore assemblages "...without presupposing that information resides only in 'end-products'" (Hiscock and Clarkson 2000). These approaches challenge archaeologists to find quantitative measurements of the rate and nature of changes to stone artefact morphology that occur during flaking. Consequently, one of the consuming methodological questions in studies of stone artefacts is to identify robust and reliable measurements of the intensity with which stone was reduced. A large number of methods have been suggested and employed, and this paper undertakes a review of prominent quantitative approaches advocated recently. In particular our focus is on one measure, the Reduction Index proposed by Kuhn (1990); a measure which has been subject to negative comments (e.g. Dibble 1995) but which has been used extensively in Australian archaeology (e.g. Clarkson 2002a, Lamb, this volume and Law, this volume; Hiscock and Attenbrow 2002; 2003, this volume). In this paper, we use experimental data to provide a quantitative description of the relationship between the index and the rate of change to retouched flakes during reduction. This experimental evidence supplies the basis for a revised comparison of the different methods of measuring the intensity of retouching on retouched flakes.

Measures of Reduction Intensity

The measurements of reduction intensity that have been proposed are diverse. Approaches broadly fall into four categories: 1. analysis of the relative abundance of different implement classes within an assemblage, 2. description of the nature of the retouching, 3. estimation of the original blank size, and 4. quantification of the extent of retouch scars. In this section we characterize and review a number of approaches.

Relative Abundance of Different Implement Classes

One approach to examining the extent of reduction displayed by an assemblage is to compare the frequency of specimens in different implement types. By hypothesizing that some implement types result from minimal reduction while other types have been heavily reduced, it becomes possible to interpret the proportions of implement types as an indicator of the typical intensity of reduction represented in an assemblage. An outstanding example of this kind of interpretation is Dibble's (1988:189) use of a 'scraper reduction index' to express the numerical abundance of convergent, transverse scrapers and Mousterian points relative to single and double scrapers as a way of inferring the emphasis on different kinds of reduction in European sites. Similar approaches to expressing the extent of reduction have been applied to Australian assemblages. One example is Hiscock's (1994) hypothesis that north Australian bifacial points had been more heavily worked than unifacial points and that consequently the percentage of points that are bifacial (or the bifacial:unifacial ratio) could be employed as a measure of the amount of reduction in an assemblage. Of course the efficacy of such analyses is largely dependent on the accuracy of the assertions about the position of each type in the reduction process, and it is partly to this end that researchers have searched for generic and robust measures of reduction suitable for analysis of individual specimens.

A variant of this approach is the observation of the abundance of shaping and resharpening debris relative to implements. In analysing knapping technologies, in which extending reduction creates more flakes, many archaeologists have therefore employed the number of flakes per 'tool' as an expression of the extent of retouching. The flake:tool ratio should be higher when the average extent of reduction is higher. The same argument has been given in analyses contrasting Australian assemblages (e.g. Hiscock and Allen 2000). These ratios typically assume that artefacts discarded on a

site were knapped there, and are reliable expressions of the average number of flakes struck. Such assertions can only be made, however, when the net loss or gain accruing from artefact transportation between sites is minimal, a proposition that is difficult to test under most conditions.

Nature of Retouching

A number of archaeologists have examined the nature of retouch scars as an indication of the extent of reduction. In the context of Palaeolithic Europe this has most often been accomplished by analysing the invasiveness of retouch scars, but the angle of retouched edges and the state of flake terminations are also traits that have been cited. These analyses have involved a comparison of the typical retouch characteristics between each implement type. For example, Dibble (1984:433) recorded specimens using a four state ordinal variable he named 'retouch intensity' (light, moderate, heavy or stepped) and argued that implement types with the highest frequency of higher retouched states were those that had been most intensively worked. Dibble (1984:434) summarised his inference by concluding:

Assuming that the level of retouch intensity corresponds in part to the amount of material removed during retouching, then these data suggest that there is an increase in the amount of modification as one moves from the single, through double, to convergent scrapers.

Gordon (1993:209) used a similar system to that of Dibble for analysis of flake reduction at the Mousterian site of Ghar in Israel. Gordon's system comprised five ordinal rankings between 0 for no retouch and 4 for retouch formed of more than two rows with deep wide scars. The reliability of ranking systems such as these depends on a number of factors, including the consistency of the classification, the accuracy of the ordinal rankings as a measurement of the extent of reduction, and the discreteness of the typological classes. Furthermore, the directionality of changes through the retouching process in traits such as edge angle, scar size and scar termination have not been independently established, either experimentally or through inspection of archaeological materials, and the dependability of ordinal categories as a measure of retouch intensity remains unclear.

Estimating the Original Blank Size or Mass

A third and very extensively exploited approach to measuring the amount of reduction is to estimate the size or mass of the flake prior to retouching. These estimates may then be employed to calculate the amount of material that has been removed during retouch. This approach has been emphasized by several researchers as one of the better indications of the intensity of reduction (Dibble 1997; Dibble and Pelcin 1995; Holdaway 1991). Its application is based on two propositions.

Firstly, that the original size or mass of a flake can be estimated from a number of attributes. These include

using measurements of thickness and platform area as a means of calculating original flake size. Platform features are regarded as critical because they often remained intact while lateral and distal portions of the flake were retouched: hence platform features can be measured and used to estimate ventral surface area. Regression analyses between thickness or platform area and the ventral surface area of unretouched flakes in archaeological assemblages serve as the empirical basis for predictive statements. An extensive list of correlation coefficients is provided by Dibble (1995:326), for a large number of Palaeolithic assemblages. Almost all are statistically significant at the $p = 0.005$ level, giving him confidence in the predictive ability of this measure. The predictive capacity of platform dimensions (platform thickness, platform width and external platform angle) as an estimator of original flake mass, on the other hand, has been examined in a number of studies using the controlled experimental fracture of simple glass cores (Dibble and Pelcin 1995; Dibble and Whittaker 1981; Pelcin 1997a, b, c, 1998) and from archaeological and experimental assemblages (Dibble 1997; Shott *et al.* 2000).

The second proposition required to transform these attributes into measures of retouch intensity is that the estimates of original size or mass can then be used to calculate the amount of stone lost from a flake through retouching. For instance, it is argued that the ratio of platform area to ventral area (Dibble 1995), or of thickness to ventral area (Holdaway *et al.* 1996), can give an indication of the amount of surface area lost from a flake through retouching. Original size or mass estimated from platform characteristics, on the other hand, can be compared to the observed mass of a flake to express the amount of stone lost through reduction (Dibble and Pelcin 1995). While Dibble advocates undertaking these analyses at the assemblage level to give an indication of the average level of retouching intensity in that assemblage (Dibble 1987b:113; 1997, 1998), others see potential to develop predictions of original mass that will accurately measure retouch for individual specimens (Pelcin 1998; Shott *et al.* 2000).

Dibble (1995:327) argued that "because of its ability to help control for original blank size, the ratio of surface area to platform area is an important variable in demonstrating scraper reduction". The evidence Dibble cited undoubtedly shows that within individual assemblages that have been created with a limited range of technological strategies the correlation of ventral area with platform area and thickness enables estimates the typical size of each implement class before retouching began. This has been a valuable inference in his quest to understand the relationship of the Bordesian types to each other. However, the usefulness of this measure is limited by the generally low explanatory capacity of these correlations.

Other indications that predictions of original size may be

unreliable have come from more recent experimental studies. Pelcin (1997b), for example, found that for a given platform area the resulting ventral area of a flake will vary according to indenter type, and that surface area therefore cannot be accurately predicted from platform attributes when indenter type is unknown (which is usually the case in archaeological assemblages). In Shott *et al.*'s (2000:888) analysis of experimentally knapped assemblages, ventral area did not correlate to platform area as well as flake mass. They partly attribute greater variation in platform area correlations to the error introduced through the use of imprecise measures of ventral area (e.g. length x width), but acknowledge that core form may also have an effect on surface area independent of platform size. Kuhn (1990) has also observed that platform size generally accounts for no more than about 20% of the variation in surface area of unmodified flakes in Mousterian assemblages he has examined.

Davis and Shea (1998), on the other hand, attempted to evaluate the performance of Dibble and Pelcin's (1995) predictor equation of original flake mass using platform thickness and external platform angle. Not surprisingly given the shift from controlled laboratory conditions to uncontrolled knapping procedures, Davis and Shea found that predicted weight deviated from actual weight by at least $\pm 10\%$ in most cases, but by as much as 175% in others. They concluded that caution must be shown in applying this predictor in its current form to archaeological assemblages due to its generally poor performance, and pointed to the omission of platform width as a likely source of much of this error.

In his reply to Davis and Shea, Dibble (1998) concurred on the issue of platform width, but also emphasized the added variation introduced by the greater complexity of real-life knapping situations where many variables are allowed to vary freely. In Pelcin's (1998) separate reply to Davis and Shea, he disagreed with both Dibble and Davis and Shea over the matter of platform width, arguing instead that this variable only had a threshold effect in determining flake mass. Pelcin saw modeling of knapping patterns and raw materials for individual assemblages as the best way to proceed from controlled fracture of glass cores to real assemblages. Shott *et al.*'s (2000) analysis of experimentally knapped assemblages came to a similar conclusion regarding platform width, stating that "platform width's influence on flake size seems limited". They also found that "the relationship of mass to platform dimensions is even more variable in assemblages than in individual flakes", contradicting Dibble's assertion that whole assemblages represent the most appropriate scale of analysis. They concluded that while predicting original mass for individual specimens was beyond the ability of current methods, it is still worth continuing efforts to do so.

Hence the principle of inferring the extent of reduction by employing platform features to reconstruct original flake

mass is theoretically sound and remains the focus of ongoing investigations. However a great deal of uncertainty still surrounds the level of precision achievable in predictions of original flake size, the most appropriate scale of analysis, and whether accurate prediction is in fact achievable at all in archaeological contexts. These methodological complexities have encouraged some researchers to explore other approaches, particularly models measuring the dimensions of retouch scars.

Quantification of the Extent of Retouch Scars

The fourth approach to measuring reduction intensity is to observe the size and abundance of retouch scars on flakes. Perhaps the simplest variant is the measurement of the length of flake margins that were retouched (e.g. Barton 1988). While that trait will be of use for many analytical purposes, it is unclear how length of marginal retouch is associated with absolute expressions of reduction such as investment of time or effort, or the loss of original mass or volume. Hence the calibration of various scar measurements with absolute measures of reduction is critical. Furthermore, the results of these indices are not strictly comparable, between specimens or assemblages, due to variation in original flake size. Such measures could be considerably improved by calculating retouch extent as a percentage of edge length or flake width, rather than an absolute measure. Such an approach was employed by McPherron and Dibble (1999) using digital image analysis and has subsequently been employed by other researchers (e.g. Hiscock and Attenbrow 2003).

Marcy (1993) took a different approach to measuring retouch coverage, using digital image analysis to calculate the proportion of surface area covered by retouch. Yvorra (2000) also employed this technique but added measurements of retouched edge angle to differentiate between steep and marginal, and low-angled and invasive, retouching. It would appear that image analysis techniques such as these offer very accurate measures of retouch coverage; however, they tend to be slow and expensive and as yet few analysts use them on a regular basis.

A different procedure for assessing scar abundance is Clarkson's (2002b) estimation of retouch scar coverage. His 'Index of Invasiveness' calculates intensity of retouch by estimating the extent of retouching around the perimeter of a flake as well as the degree to which it encroaches onto the dorsal and ventral surfaces. The index is calculated by conceptually dividing an artefact into eight segments on each face. Each segment is then further divided into an inner 'invasive' zone, ascribed a score of 1, and an outer 'marginal' zone, ascribed a score of 0.5. Scores of 0 (no retouch), 0.5 (marginal) or 1 (invasive) are allocated to each segment according to the maximum encroachment of scars into one or other of these zones. The segment scores are then totaled and divided by 16 to give an index between 0 and 1. Clarkson

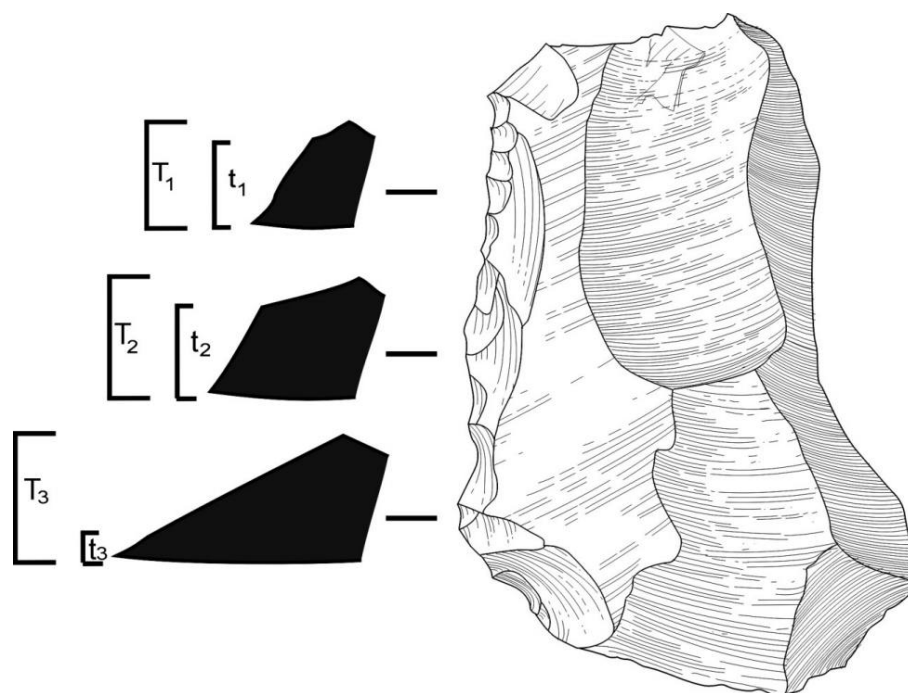


Figure 1. Illustration of the measurement of (Kuhn 1990) Geometric Index of Unifacial Reduction on a unifacially retouched flake using Method A.

(2002a: 68-71) provided experimental evidence for a strong and significant positive relationship between the index and the number of retouch blows ($r^2 = 0.982$, $p = <0.001$) and the percentage of original weight lost from each specimen ($r^2 = 0.968$, $p = <0.001$). The rate of increase for the index of invasiveness is slightly curvilinear when plotted against both independent measures of reduction, but can be made linear using a square root transformation of index values. Little variation is evident in the rates of index increase between raw materials of varying fracture quality. The index of invasiveness has the advantage of being fast to calculate and versatile, and is well suited to the measurement of both unifacial and bifacial retouch with minimal inter-observer error (Clarkson 2002a: 71).

A limitation of techniques measuring the extent of retouch on a surface is that they are less suited to assemblages in which artefacts exhibit predominantly steep and marginal unifacial retouch, as might commonly occur on backed artifacts or steeply retouched scrapers (e.g. Quina type retouch). For instance, the index of invasiveness would not readily increase above 0.25 in such cases, no matter how much reduction takes place. In assemblages with non-invasive marginal retouch, alternative measures of reduction may be more appropriate, such as Kuhn's index of reduction.

Kuhn's Index of Reduction

A measure ideally suited to estimating the amount of reduction on marginally and unifacially retouched flakes was proposed by Kuhn (1990). The index calculates the extent of retouch by the relative 'height' (ventral-dorsal)

of retouch scars. Kuhn presented two different methods for calculating what he named the Geometric Index of Unifacial Reduction, but which we refer to here as the Kuhn Index.

The first method calculates a quantitative measure of edge attrition by dividing the height of retouch scars above the ventral face ("t") by the maximum thickness of the flake ("T"). Both measurements were taken at right angles to the ventral surface and at the same point on the retouched edge (Figure 1). Both "t" and "T" can be measured directly using calipers, a technique which we will refer to as Method A.

Kuhn suggests the use of a second, more complex method to overcome problems in accurately determining the true height of "t" given variation in the curvature of the ventral face. We term this second calculation Method B. This method arrives at "t" by multiplying the length of retouch scars ("D") by the sine of the retouch angle ("a") (Figure 2). The resulting value of "t" is then divided by "T" measured with calipers, to create the index. While Kuhn argues that Method B provides more precise and replicable results, it is difficult to see how edge angle can be measured any more accurately than the height of retouch when the ventral surface is curved.

The index calculated using either method ranges from 0 to 1. A value of 0 represents no retouch and a value of 1 indicates that retouch scars have intersected with, or crossed, the point of maximum thickness. Kuhn's index provides a straight forward and relatively simple way of measuring the amount of edge lost from a retouched

flake. The nature of the index means that it is not restricted to a particular shape of retouched edge and it potentially offers a versatile measure for a wide range of assemblage types. However the index has been criticized on a number of grounds. One limitation that was acknowledged by Kuhn (1990) was that the index could only be measured on unifacially retouched flakes on which blows were applied to the ventral face and created scars on the dorsal face. Because both t and T are oriented to and measured from the ventral face, any retouching onto the ventral surface will make calculation of a Kuhn Index at that point impossible. Consequently where ventral and dorsal retouch exists on different edges of a single specimen the Kuhn Index will express the amount of retouch on only some edges. Furthermore, unifacial implements with ventral retouch and bifacially flaked specimens cannot have a Kuhn Index calculated. This restricts the proportion of an implement assemblage that can be assessed using the index, although in many parts of the world dorsally flaked uniface are the dominant category of implement. Regions in which implements are typically bifaces may have limited use for the index.

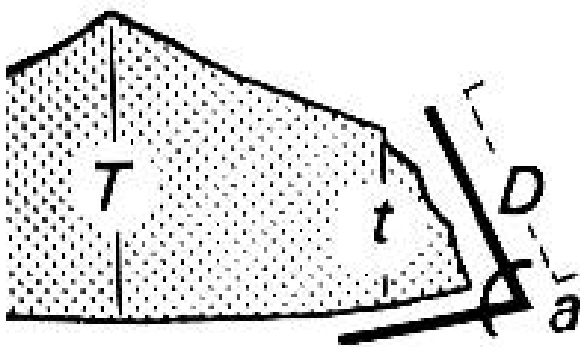


Figure 2. Illustration of the measurement of Geometric Index of Unifacial Reduction on a unifacially retouched flake using Method B (Kuhn 1990).

An additional complexity of the Kuhn Index is that retouch located at the distal end of a flake may be less altered by reworking than retouch positioned on the lateral margins. This occurs because it may take less retouch to attain the maximum value of 1 at the distal end than on a lateral margin. For this reason some archaeologists argue that the Kuhn index is only viable on laterally retouched implements. We suggest such a position is an over-reaction to the effects of cross-sectional shape, and we will return to this issue later in the paper. For the moment the significant point is that this complication with distal end measurement is in fact a special form of what we call the 'flat-flake problem'.

The most extensively developed critique of the Kuhn Index was provided by Dibble (1995:330), who argued that while the index functioned as designed on flakes with triangular cross-sections it was unresponsive to

retouching on flakes with flat dorsal surfaces parallel to the ventral face. Using the illustration we reproduce in Figure 3, Dibble explained this 'flat-flake problem' as follows:

A problem occurs in the case of very flat flakes, however, where this ratio will approach the maximum much more quickly (i.e., after fewer reshaping episodes) than it will on more highly convex flakes... Thus, while Kuhn's Reduction Index can reflect the amount of retouch that is applied, it will also be affected by the exterior morphology of the flake. Though more objective than the previous technique, it is still not an unambiguous measure of how much material was removed.

The theoretical point that the rate at which the index changes is probably related to flake cross-sectional shape is, we argue, correct, and an appreciation of that effect should be built into interpretations of the index. However, the magnitude of this effect has not been empirically measured and its impact on the interpretation of retouch intensity using Kuhn's reduction index has not been established. While Dibble's critique is technically correct, it has not been shown to create a significant problem for interpretation of most archaeological assemblages. To assist in evaluating the robustness of the index, and examine the likely impact of the 'flat-flake problem' we proceed to a re-evaluation and experimental testing of the Kuhn index.

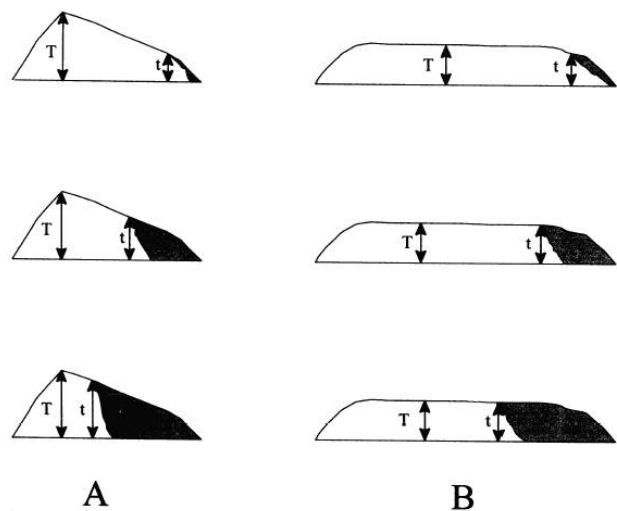


Figure 3. Dibble's (1995:329) illustration of the 'flat-flake problem'.

Evaluating Kuhn's Index

The Kuhn Index can be reliably repeated but its interpretation must be informed by a number of considerations. There are three questions that must be answered in order to interpret the index:

1. Is the index invariably positively correlated with the intensity of reduction?,

2. Is that correlation linear or non-linear?, and
3. In what conditions do those patterns vary.

Our consideration of these three questions begins with the arguments advanced by Kuhn in his initial discussion of the index.

As a way to evaluate the effectiveness of this index, Kuhn (1990) performed a series of experiments involving the retouching of 22 flakes. Each flake had an edge flaked on a number of occasions, called 'events', to simulate maintenance of a working edge. At the completion of each retouching event Kuhn measured the reduction index in two ways: by a single observation in the centre of the retouched edge (called the "centre edge" value) and by the mean of three observations along its length (called the "mean" value). On the basis of these experiments Kuhn (1990) was able to derive a number of inferences:

- Both forms of measurement reveal that the index values increase as the number of retouching events increases, so that there is a positive relationship between number of events and size of the index.
- The values of centre-edge and mean values typically differ, with the centre-edge index often being higher.
- There is considerable variation in the amount of change to the index that occurs between retouching events.
- Kuhn suggested that the relationship between retouch event and reduction index was slightly curvilinear.

A re-analysis of Kuhn's (1990) published experimental results reveals a number of further points, and a revision of his conclusions. Firstly, centre-edge measurements will often display a flatter curve with a larger range of values than mean measurements, even though the central tendencies are nearly the same for both measurement systems ($\bar{X} = 0.55$ for centre-edge and $\bar{X} = 0.53$ for

mean, $N=118$). This occurs because the averaging effect doesn't simply lower mean values relative to centre-edge ones it also concentrates values around the central tendency, making the distribution of mean values display more pronounced kurtosis (see Figure 4). The consequence is that in Kuhn's experiments centre-edge measure values often ranged up to 1 but low values (less than 0.1 - 0.2) were rare; whereas with the mean measure both high (0.85-1) and low values (less than 0.1 - 0.2) were rare.

This observation implies that the relationship between centre-edge and mean values is not adequately depicted simply as centre-edge values being larger. As shown in Figure 4 a linear regression between paired centre-edge and mean values shows that the two values are strongly correlated ($r^2 = 0.945$).

Figure 4 also illustrates Kuhn's observation that there is noticeable variation in the amount of change to the index that occurs between retouching events. The overlap between the reduction index values returned within retouching events 4-6 is greater than the overlap between events 1, 2, and 3. This decrease in the magnitude of change in index values per event as reduction proceeds, from about 0.14 early in the sequence to only 0.4 - 0.7 later in the sequence, was the subject of extended analysis by Kuhn (1990) and will not be elaborated here, although we return to this point in the discussion of our own experiments below.

Change in the magnitude of index increase between retouch events underlays Kuhn's conclusion that the relationship between retouch event and reduction index was slightly curvilinear. While that may be true, our reanalysis of his experimental data suggests that Kuhn may have over-emphasised the non-linearity of the relationship. Linear regressions of both mean / event and centre-edge / event pairs show impressively high

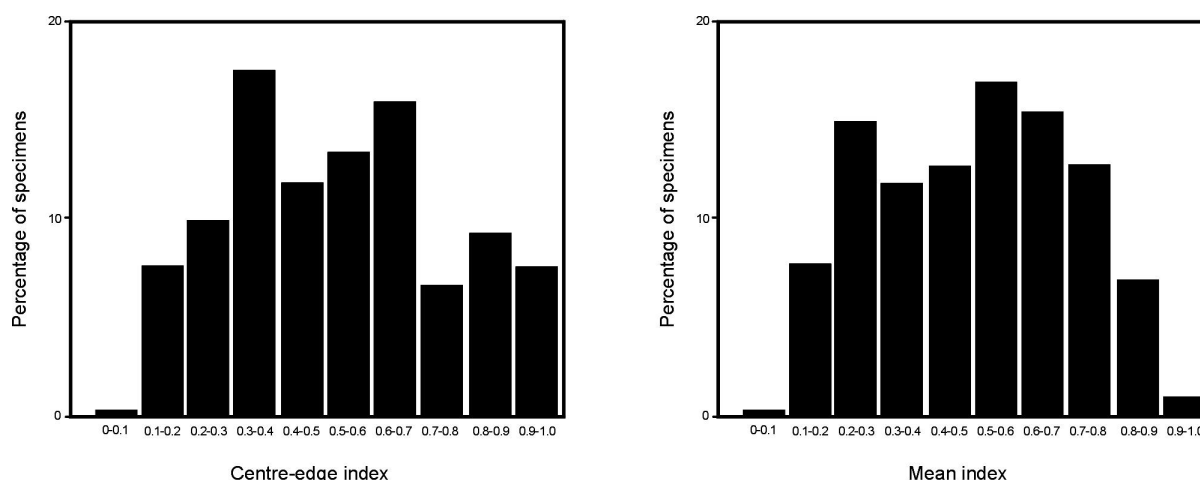


Figure 4. Histograms showing the values of the centre-edge and mean indices produced during Kuhn's (1990) experiments.

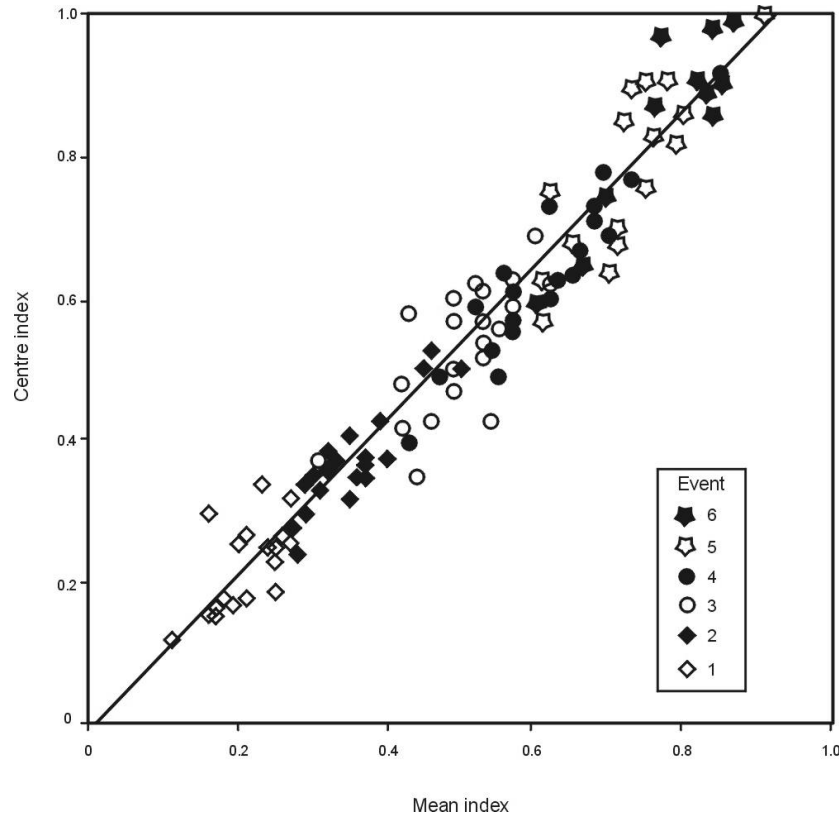


Figure 5. Scattergram comparing the values of paired centre-edge and mean measurements produced during Kuhn's (1990) experiments.

Table 1. Relationship between the variables in Kuhn's (1990) experiment as expressed by Pearson's coefficient and (Spearman's coefficient).

	Centre-edge	Mean	Event
Centre-edge	1.000 (1.000)	0.972 (0.974)	0.902 (0.909)
Mean	0.972 (0.974)	1.000 (1.000)	0.929 (0.936)
Event	0.902 (0.909)	0.929 (0.936)	1.000 (1.000)

All correlations are significant at $p = 0.001$ for both test statistics. $N = 118$ for all pairs.

coefficients using either Pearson's product-moment statistic or Spearman's (Table 1). These coefficients reveal that for Kuhn's experiments the number of retouching events explains more than 80% of the variation in centre-edge values ($r^2 = 0.814$) and more than 85% of mean values ($r^2 = 0.863$) - a conclusion almost identical with that of Kuhn (1990:591). This connection between the extent of retouch and Kuhn's reduction index is impressive and given suitable flake morphologies should give analysts confidence in inferring the relative intensity of retouching from either of the Kuhn indices - especially in contrast to the low predictive power of area/platform area indices discussed above.

Our main concern about Kuhn's experiments is his use of the 'retouching event' to measure reduction. Despite the care that he took in conducting the experiments, Kuhn's

choice of this unit of observation was a poor one since there is no reason to believe that these events were of equivalent magnitude to each other; either within or between experimental specimens. Hence, while we accept that Kuhn's experiments demonstrate that the reduction index displays a unidirectional relationship with the extent of reduction, we do not accept his experiments as an adequate demonstration of the linear or non-linear nature of the relationship. Despite the high linear correlations displayed in the experiments it is possible that the use of retouching events has either created the impression of a curvilinear relationship where a very linear one exists or, alternatively, has created the impression of a strong linear relationship while hiding the non-linear nature of the relationship. We believe that an exploration of the linearity of the relationship between the extent of reduction and Kuhn's reduction index

should be conducted using weight of rock removed and/or number of flakes removed during retouching. To this end we conducted an experiment that was very similar to Kuhn's but in which we measured changes to mass as well as numbers of flakes struck.

An Experimental Re-Evaluation

Methods

The methods chosen to evaluate Kuhn's index are similar to those undertaken by Clarkson (2002b), and involved tracking changes in the rate of increase in index values against numbers of retouch blows and the percentage of weight lost from each specimen. By establishing the nature of the relationships between these variables, we hope to determine the degree of linearity, the actual as opposed to theoretical range of the index, and the limitations of this approach for measuring retouch.

The experiments involved unifacial percussion flaking of thirty flakes. Blows were applied to the ventral face of one lateral margin, removing flakes from the dorsal face to create a straight retouched edge. This was done in a number of episodes, each comprising ten flake removals more than 3mm in length positioned along the entire length of the specimen, and at the end of each retouching episode a number of attributes were recorded on each specimen. This provided a record of the progressive changes in morphology for each specimen during reduction, and gave a total of 348 data points. The approach to reduction was conservative, with the authors aiming to remove enough of the edge to effectively resharpen or rejuvenate it, but without removing unnecessary mass. To avoid judgments on functionality, retouching was continued until the specimen broke.

A summary of the experimental results is given in Table 2. The amount of reduction varied, with as little as 68 flakes and as many as 203 flakes being removed before specimens broke. This resulted in an average weight loss of approximately half the original weight of flakes,

although the percentage of weight removed varied between specimens. All specimens had attained high Kuhn reduction index values before they were broken.

This experiment held many factors constant, including raw material (mudstone), the technique of retouching (direct hand held percussion), the face retouched (dorsal), the number of margins retouched (one), the shape of the retouched edge (straight), the interval between measurement (10 blows), and the weight of hammer stones (two hammers weighing 82gm and 55gm were used throughout). The main factor that was varied was the flake blank, as a way of evaluating the effect of flake morphology on the development of high values of the Kuhn reduction index. We created a number of flakes that were broadly similar in size to those retouched in prehistoric Australian assemblages. As summarized in Table 3, these flakes were quite varied in weight (27-344g), width (29-89mm), thickness (8-33mm), cross-section (steep triangle to flattish trapeze, see Figure 6), number of ridges (1-4), and edge angles (32°-104°). We intend to explore the relationship of these aspects of flake morphology to changing values of the Kuhn reduction index on another occasion; here our only purpose is to evaluate those trends in the Kuhn index that are so robust they exist despite this massive variation in blank morphology.

Results

In this experiment the number of blows has a complex relationship with the Kuhn reduction index. As shown in Figure 7, the experimental data points display a wedge-like pattern with low reduction index values having been reached in only a few blows but high index values being associated with both large and small numbers of flakes, reflecting wide differences in the number of flake removals required to achieve large Kuhn values. While the correlation is statistically significant the coefficient reveals that the relationship is only moderately strong ($r = 0.716$, $r^2 = 0.513$, $r_s = 0.748$, $N=348$, $p<0.001$). The

Table 2. Summary of experimental results

	N	Minimum	Maximum	Mean	$\pm$	Std. Dev.
Number of flakes	30	68	203	111.90	$\pm$	31.73
Kuhn reduction index	30	0.79	1.00	0.95	$\pm$	0.06
Percentage weight loss	30	15.0	82.3	51.73	$\pm$	16.65

Table 3. Summary of experimental flake blanks

	N	Minimum	Maximum	Mean	$\pm$	Std. Dev.
Weight	30	27.3	344.2	75.11	$\pm$	67.37
Length	30	49.2	119.7	72.45	$\pm$	15.82
Width	30	28.9	88.5	45.74	$\pm$	13.82
Thickness	30	8.0	32.7	16.11	$\pm$	6.46
Number of ridges	30	1	4	1.93	$\pm$	0.58
Average edge angle	30	32.3	103.7	51.5	$\pm$	15.6

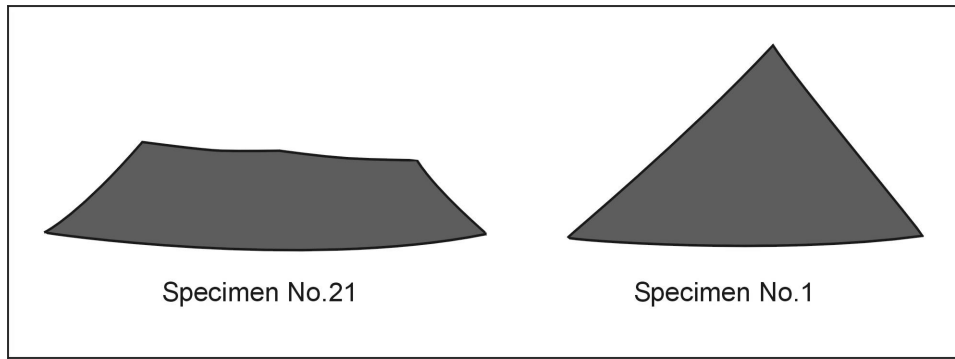


Figure 6: Examples of the range of flake cross-sections used in the reduction experiment. Specimen No.21 has a flattish trapezoidal-like cross-section similar to Dibble's 'flat flakes', while Specimen No.1 has a steep triangular cross-section.

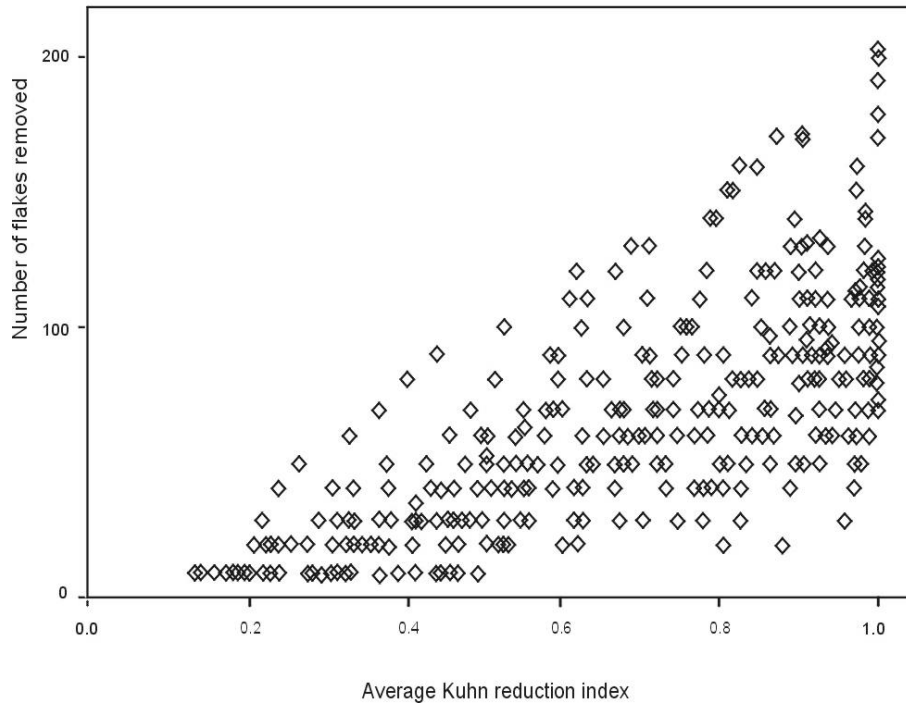


Figure 7. Scattergram depicting the relationship between the number of flakes removed and Kuhn's Index of reduction in our experiments.

primary cause of this pattern is the variation between flakes in mass removed. A more robust description of the relationship of reduction and the Kuhn index is achievable by focusing on mass removed.

A different depiction of the relationship between the extent of reduction and the Kuhn index is found when the percentage of weight of the original flake that has been lost during retouching is used as the measure of reduction. Figure 8a plots percentage of weight lost against the Kuhn index for our experimental specimens. The datum points show that there is a discernable and strong positive relationship between the mass removed during retouching and the Kuhn index values generated

by that retouch. This graph of experimental data also reveals that weight loss is related to the Kuhn index in a distinctly non-linear manner. The covariation is approximately log-linear in nature; a pattern that is clear in Figure 8b, which reduces the data to a series of bars displaying the 95% confidence interval for the mean of each 0.1 unit of the Kuhn index. It is worth noting that every bar is separated from and lies entirely above the preceding one – revealing the strength of positive covariation. This depiction of the trend makes the log-linear relationship between the variables apparent: low Kuhn index values are attained by removing a small amount of material whereas on extensively retouched specimens the removal of a proportionately large amount

of material produces only small changes in the Kuhn index.

There are a number of reasons for the non-linear nature of this association. Flake geometry is partly responsible. On many flakes the increase in thickness away from the lateral margin means that similar blows will remove less mass from the margins of the flake, early in the retouching process, than from the centre of the flake, later in the process. The nature of reduction also changes as retouching continues, with the creation of steep angles and step terminated scars compelling the knapper to rejuvenate the edge by striking bigger and more invasive flakes, creating longer scars. Furthermore, since the Kuhn index, by definition, has a maximum value of 1 and reduction can continue after that value is reached, the

relationship must become non-linear as retouching continues, because on heavily retouched specimens mass is lost without altering the Kuhn index.

The curvilinear relationship of the Kuhn index to mass reduction is significant for interpretations of the index. Since relatively more weight is lost later in the flaking sequence than early in the retouching process, not all increments in the Kuhn index are equivalent. For example, in terms of mass lost the interval between 0.8 and 0.9 is substantially greater than between 0.2 and 0.3. Consequently, comparisons between assemblages and sections of assemblages that have different values of the Kuhn index should be couched in terms of relative rather than absolute differences in the extent of retouch, unless a relevant calibration is available.

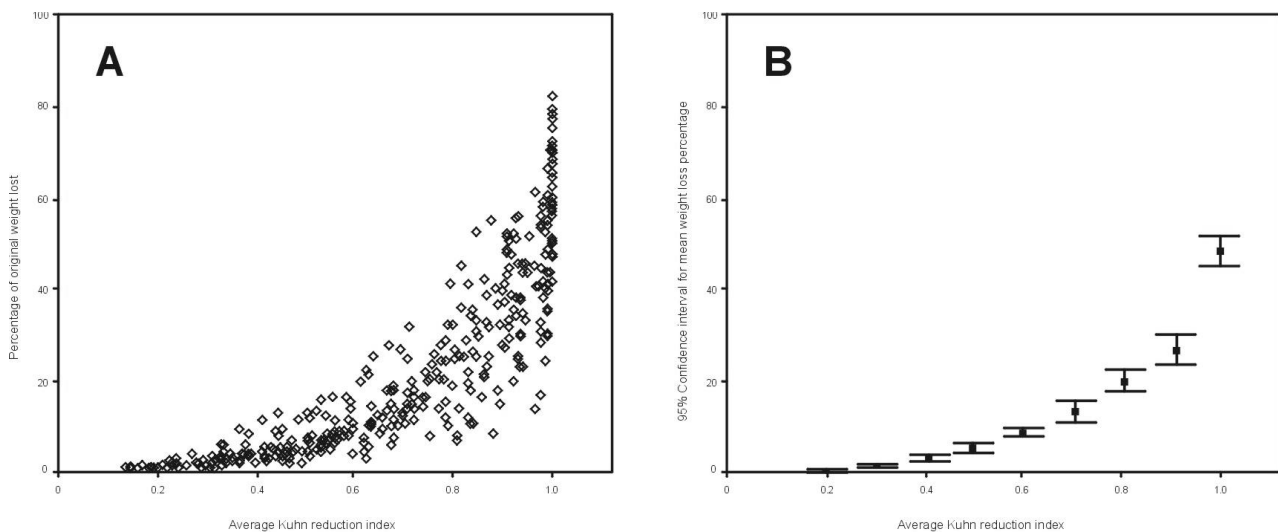


Figure 8. Illustrations of the relationship between the Percentage of original mass lost and Kuhn's Index of reduction for our experimental specimens. A is a scattergram of the raw data. B shows bars displaying the 95% confidence interval for the mean of each 0.1 of the Kuhn index.

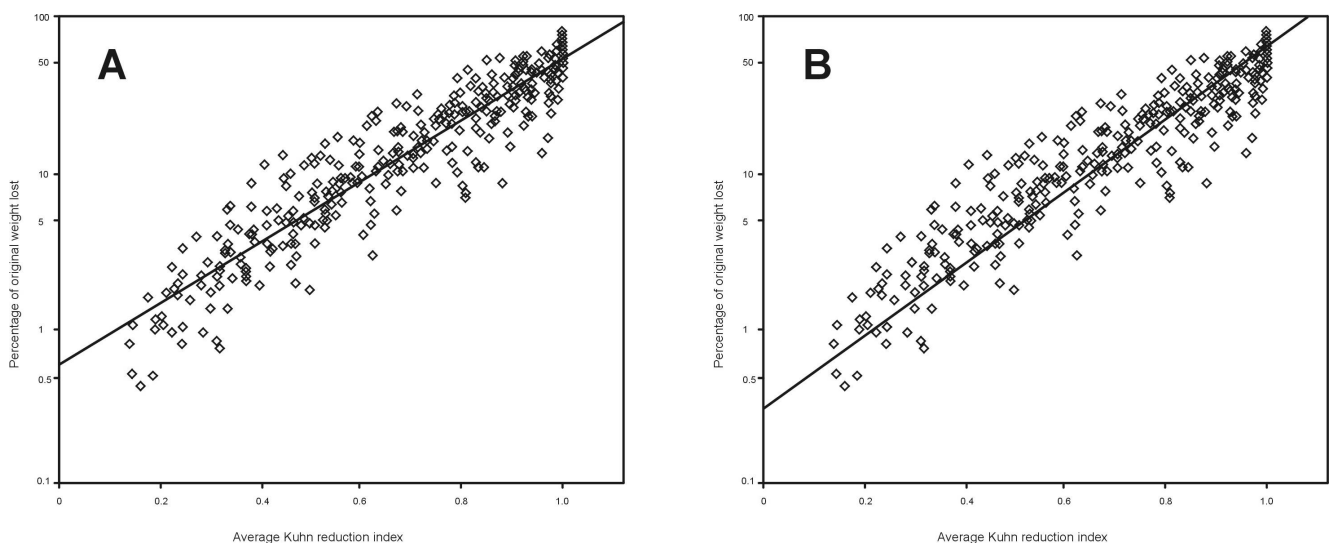


Figure 9. Illustrations of the relationship between the percentage of original mass lost expressed on a log scale and Kuhn's Index of reduction for our experimental specimens. A Shows the regression line calculated with a constant ($r^2 = 0.871$), while B shows the line of best fit constructed without a constant ($r^2 = 0.985$).

Depictions of the data provided in Figure 8 reveal several further patterns. Firstly, the minimum value recorded for the Kuhn reduction index, on specimens with minimal retouch, was 0.14. This demonstrates that even in the initial phase of retouching values less than 0.2 may be rare, and depending on the definition of retouch employed, values less than 0.1 may not be found in many assemblages; a pattern congruent with the results of Kuhn's (1990) own experiments.

A second observation is the continued loss of mass through retouching on some specimens after a value of 1 has been reached. Twelve specimens, 40% of the experiments, reached Kuhn values of 1 before breaking. Those specimens reaching values of 1 did so when weight loss was 57.1 ± 8.3 percent of the original flake ($N=12$). For those specimens 13.1 ± 7.7 % of the original flake weight was removed after values of 1 were recorded. It should be emphasized that the conditions of our experiments exaggerate this effect, because all specimens were reduced until they were broken.

The implications of these findings are:

1. Although in theory the index is scaled from 0, in practice the range of values will usually be less, starting between 0.1 and 0.2,
2. While the maximum value of Kuhn index is typically reached when 50-65% of original mass has been removed, specimens with values of 1 represent varying levels of reduction and should not necessarily be interpreted as a maximum or near maximum amount of retouch, and
3. In relation to the change in the relative mass of each flake produced by retouching the Kuhn reduction index is not linearly scaled and should not be interpreted as though it was. The reduction index can reliably be used as a relative measure of the amount of mass removed, but a further analytical step is required to 'calibrate' it and allow it to be used as an absolute measure.

Our experiments indicate that in some instances the transformation of variables may be sufficient to create a strong linear relationship, thereby providing a basis for absolute statements of different levels of reduction. For our experimental data it is a simple matter to re-express the percentage of original flake weight lost through retouch on a logarithmic scale, thereby transforming the relationship of mass loss and the Kuhn index into a linear one. The scatterplot resulting from this transformation is illustrated in Figure 9. A linear regression of these data, calculated with a constant, gives a correlation coefficient of 0.933 ($N=348$, $p<0.001$), which can be interpreted as 87% of the variation in mass loss being expressed by values of the Kuhn index ($r^2 = 0.871$). A similar analysis, without constant, gives a coefficient of 0.993 ($N=348$, $p<0.001$), a remarkably high value that indicates that approximately 98% of mass loss is explicable in terms of the Kuhn reduction index ($r^2 = 0.985$). With correlations

coefficients of these strengths it is reasonable to assert that, at least in single margin reduction of the type experimentally tested, the percentage of weight lost could be reliably predicted from the value of the Kuhn reduction index that can be measured on specimens.

Kuhn as a Predictor of Extent of Reduction

The experiments we have described here indicate that the Kuhn reduction index is a poor predictor of the number of flakes removed, but is a robust indicator of the progressive loss of weight from a retouched flake worked on a single lateral margin. The relationship between loss of mass and the reduction index is non-linear, with relatively more weight lost later in the retouching process per measured interval. This pattern must be considered in deriving interpretations based on the Kuhn reduction index, and we suggest that inferences can be based on the principle that the value of the index measures $\log(\% \text{weight loss})$. Treated in this way the Kuhn index is a reliable description of the amount of flake retouching. We particularly note that the flakes we retouched were selected to represent a large variety of cross-sections, ranging from very flat to steeply triangular. The strong non-linear correlation displayed by our experimental data therefore provide grounds for concluding that the flat-flake problem discussed by Dibble may exist but need not create an obstacle to employing the Kuhn reduction index as a powerful way of measuring the extent of flake reduction.

It remains to be seen how well the Kuhn index performs as a measure of extent of reduction outside of the parameters set for this experiment. Retouching one, two or three additional margins, for instance, or adopting patterns of retouch that begin on one margin and expand outwards versus those that begin on separate margins and converge toward a point may perhaps create quite different index to mass relationships to those documented here. It is for future experimentation to resolve this issue.

A further consideration is how well the index performs in the measurement of distal retouch. A theoretical expectation at least is that the measurement of distal retouch should present difficulties for the consistent measurement of reduction due to variation in cross-section shape found along the percussion axis of flakes. Flakes that taper little over their length for instance would lead to very little increase in the index in the same way that Dibble predicted for flat flakes. However, we believe that the suitability of the Kuhn index to the measurement of distal retouch is not out of the question, but merely involves careful attention to determining the types of blanks that might be suited to this kind of analysis, enhanced by experimental studies designed to evaluate various cross-sectional shapes similar to those conducted here.

For instance, a common form of retouched flake found in arid regions of Australia is known as the *tula*. Specimens of this type exhibit a very large bulb of force with the

point of maximum thickness often near the proximal end and retouching is unifacial and typically only at the distal end. Because the thickest part of a tula is normally at or near the platform the Kuhn Index may only approach or attain a value of 1 when retouch scars began to remove the platform edge. It is therefore possible to find examples of distally retouched flakes that might be suited to measurement of the Kuhn index, at least once appropriate calibrations for various retouch patterns are devised. Future experiments might well resolve such issues, and ideally such studies should endeavor to produce a list of criteria for suitable flake blanks and accompanying calibrations that enable the use of the Kuhn index as a measure of extent of reduction in a wide variety of situations.

Comparing the Methods

An evaluation of Kuhn's index would not be complete without a comparison of its performance to alternative measures. To provide a basis for comparing different kinds of measurements we have calculated from our experimental data a number of the different reduction measures discussed in this paper (see Table 4). For each measure we have calculated its linear correlation with changing weight loss. Table 4 provides regression coefficients for five measures of reduction, including the Kuhn Index, determined using the percentage of weight lost from each specimen as the independent absolute measure of reduction. Where appropriate we have corrected for non-linear relationships by applying a data transformation; the last column in Table 4 indicates the type of transformation that obtains the highest coefficient for each measure. Due to the design restrictions embedded in our experimental methods, we have excluded several indices that, while no doubt of great interest to archaeologists, are inappropriate in this context. Measures of retouch distribution (i.e. the index of invasiveness, % scar coverage and % perimeter of retouch), for instance do not change during the course of reduction in our experiments because we held them constant. Those measures are therefore excluded from analysis.

To develop a ranking system that in some ways approximates those used by Dibble (1995) and Gordon

(1998), but excludes any measure of retouch distribution, we have used a ranking system that incorporates only the relevant attributes of those ranking systems; that is, edge angle, scar length and frequency of step terminated retouch. To calculate this index, the range of values recorded in each variable over the sequence of reduction was divided into four equal intervals (ranks) and assigned to each specimen for each retouching event. The mean of these three rankings was calculated for each specimen, providing an overall ranking that was regressed against log percentage of original weight lost to determine the performance of these attributes as a measure of reduction over the experimental sequence.

Calculated in this way the coefficients provided in Table 4 allow a comparative judgement of the effectiveness of different measures of reduction in the circumstances of our experiment: highly variable blank forms reduced in a standard way by unifacial retouching one lateral margin. Note that because of the large number of observations available, all tests show a decidedly non-random pattern, as measured by $p < 0.001$ in every case. These significance values cannot be employed as an indication of the relative differences in predictive strength of the different measures, and we therefore adopt the simple practice of emphasising the coefficient as the apposite means of comparing the predictive power of each measure. We have ordered the various measurements by the size of the calculated coefficients, making the order in Table 4 a rank-order list of the effectiveness of the different measures in describing the proportion of original flake weight that had been lost. The Kuhn index performs extremely well compared to other indices, and explains at least 35% more variation than other measures (as revealed in an r^2 calculation). In contrast, some indices performed very badly, such as Dibble's (1995) surface area to platform area index which explains as little as 6.7% of variation. In the kind of situation represented by our experiment, such as assemblages of side scrapers, we would strongly recommend abandoning the use of a surface area to platform area index in favour of other more powerful measures. Interestingly, the variant of this index devised by Holdaway *et al* (1996) that uses thickness rather than platform area as the estimator of original flake size is far superior, explaining 47% more of

Table 4. Comparison of regression correlation coefficients for various measures of reduction.

<i>Measure</i>	<i>Coefficient (r)</i>	<i>Probability</i>	<i>Transformation</i>
Kuhn Index: Method A	0.933	<0.001	Log(% weight lost)
Kuhn Index: Method B	0.912	<0.001	Log(% weight lost)
Surface Area/Thickness	0.727	<0.001	None
Retouch Scar Length	0.697	<0.001	Log(% weight lost)
Ranked scar characteristics	0.674	<0.001	Log(% weight lost)
Surface Area/Platform Area	0.259	<0.001	None

the variation. That conclusion is also consistent with the correlation analyses presented by Dibble (1995). Close's (1991) retouch scar length and the retouch ranking system also achieve only moderate success with both explaining less than 50% of variation.

Our conclusion is that in these circumstances the Kuhn Index is the most powerful of the measures, and should be employed as a robust indicator of the extent of reduction when retouching patterns are suited to the calculation of the index.

Note that while both of the methods used to calculate the Kuhn Index gave high correlation coefficients Kuhn's preferred method of calculating the index - Method B which employs edge angle - provides a marginally lower coefficient than the simpler ratio of t/T used throughout this paper. This is an interesting result, and suggests that user-error may be compounded by the introduction of a third measurement, especially one that is notorious for its inaccuracy (Odell 1989). The difficulty of accurately measuring edge angle from a curved ventral surface would likely make edge angle measurements particularly prone to error. We therefore recommend the use of Method A in making Kuhn index calculations.

While it could be argued that the experimental techniques adopted here constrain retouching techniques beyond what might reasonably be expected in archaeological assemblages, we see the use of rigid retouching patterns as providing an opportunity for each index to perform to the best of its ability without interference from complicating factors such as variation in flaking patterns. That they have been judged and found wanting suggests that Kuhn's index is likely the most robust measure of marginal unifacial reduction currently available, both for individual specimens and assemblage-wide comparisons.

Conclusion

All reduction indices are likely to have a number of strengths and weaknesses, and while it is worthwhile considering these on purely theoretical grounds, experimental evaluations are ultimately our most effective means of determining the relative merits and operational limits of each one. Our experimental evaluation of Kuhn's geometric index of unifacial reduction indicates a level of performance that appears to be well above any of its current competitors, at least within the stringent experimental procedures we adopted. Our results suggest that Kuhn's index approximates an absolute measure of reduction once index values are recalibrated.

As the experiments were necessarily limited to the repetitive reduction of a single straight margin, however, further experiments are required to evaluate the performance of the index under a wider range of knapping situations (e.g. when more margins are included, for distal retouch etc). It is highly encouraging, however, that within such a narrow experimental

framework in which all measures should be able to achieve their best results, Kuhn's index performs well above all the other measures evaluated. We therefore see no reason to reject the use of this index in archaeological analysis.

Our experiments have also raised issues, at least in our own minds, concerning the potential of experiments such as these to understand the morphological transformations that commonly take place on retouched flakes over the course of reduction. These include changes to edge angle, the frequency of step terminations, effects of edge rejuvenation, and breakage thresholds that result from continued reduction on a single margin. It is our intention to further explore such issues in future work.

References

- Clarkson, C. 2004 Technological Provisioning and Assemblage Variation in the Eastern Victoria River Region, Northern Australia: A Darwinian Perspective. PhD Thesis. Australian National University, Canberra.
- Dunnell, R.C. 1971 *Systematics in Prehistory*. New York: The Free Press.
- Foley, R. and Mirazon Lahr, M. 2003 On stony ground: lithic technology, human evolution, and the emergence of culture. *Evolutionary Anthropology* 12:108-22.
- Hayden, B. 1977 Stone tool function in the Western Desert. In Wright, R.V.S. (ed.) *Stone Tools as Cultural Markers: Change Evolution and Complexity*. Pp.178-88. New Jersey: Humanities Press.
- Hayden, B. 1979 *Paleolithic Reflections: Lithic Technology and Ethnographic Excavations among Australian Aborigines*. Canberra: Australian Institute of Aboriginal Studies.
- Hiscock, P. 1998 Revitalising artefact analysis. In Murray, T. (ed.) *Archaeology of Aboriginal Australia: A Reader*. Pp.257-65. St Leonards: Allen and Unwin.
- Holdaway, S. 1995 Stone artefacts and the transition. *Antiquity* 69:784-97.
- Holdaway, S. and Stern, N. 2004 *A Record in Stone: The Study of Australia's Flaked Stone Artefacts*. Melbourne: Museum Victoria and AIATSIS.
- Lyman, R.L., O'Brien, M.J. and Dunnell, R.C. 1997 *The Rise and Fall of Culture History*. New York: Plenum Press.
- White, J.P. 1977 Crude, colourless and unenterprising? In Bowdler, S. (ed.) *Sunda and Sahul*. Pp.13-30.

