SUPPLEMENTARY TEXT

On Estimating Original Mass

Mass is a common measure of size or volume. Several experimental studies expressed mass as arithmetic or power functions of flake attributes like platform size and exterior platform angle (e.g. Davis and Shea 1998; Shott et al. 2000). Experimental data usually are controlled for factors like toolstone, core size and form, technology of reduction and the mechanics of blows and fracture. Such datasets are imperfect approximations, not perfect models, of empirical assemblages.

First we attempted to estimate original mass from platform area (Dibble 1987). In previous studies (Davis and Shea 1998; Dibble and Rezek 2009; Shott et al. 2000), size measures like mass and area were subject to different degrees of error. Mass was measured directly, but area of irregular objects like flake blanks was approximated as the product of length and width. This approximation was acknowledged as a source of error (Shott et al. 2000:881) compared, for instance, to the greater precision of digital measures. Similarly, Cadieux (2013:61) reported overestimation of up to 46% by platform area estimated as length times width, Dogandžić et al. (2015:9, Table 2) similar overestimation of flake perimeter and surface from linear dimensions. Three-dimensional (3D) scanning of platforms might at once increase the precision and reduce the error of area measurement.

Accordingly, as in earlier research (Shott and Seeman 2015) we wished to test Clarkson and Hiscock’s (2011) prediction equations that estimate original flake-blank size from platform area and exterior platform angle (EPA). Obviously, selected specimens had to retain the flake blank’s platform. Like Clarkson and Hiscock, we scanned all specimens and produced three-dimensional (3D) digital models, using NextEngine’s ScanStudio™. ScanStudio gives volume and surface area of specimens. We edited 3D digital models to extract platforms as isolated surfaces, whose area ScanStudio measured directly as opposed to the past crude estimation of this quantity as the product of platform width and thickness (e.g. Shott et al. 2000)(e.g. Supplementary Figure 1).

Clarkson and Hiscock’s (2011) experiments, graciously made available, involved a range of toolstones and expressed mass of detached flakes as a function of platform size and EPA. Most platform types they defined could be reconciled to our typology, but Clarkson and Hiscock’s platform size had a considerably wider range than Nobles Pond (Fig. 5). Over 70% (130 of their 183 flakes, even excluding their smallest type, focalized-platform flakes [Clarkson and Hiscock 2011:Fig. 4b], not found at Nobles Pond) exceeded the highest Nobles Pond value, 80.9 mm2, including eight whose platform area exceeded 1,000 mm2 (>1.5 in2). Again excluding their focalized platforms, mean platform area was 296.5 mm2, compared to 26.3 mm2 for Nobles Pond.

We selectively removed several flakes of limestone and other material before analyzing Clarkson and Hiscock’s data. Then we found results very close to those that Clarkson and Hiscock (2011:1067) reported for mostly flat-platform flakes by regressing natural logarithm of mass on common logarithm of platform size (in mm2) and EPA. Both transformations roughly normalized original variables’ skewed distributions; EPA’s distribution was roughly normal so, like Clarkson and Hiscock, we did not transform it. Regressing natural logarithm of mass (ln-mass) upon natural logarithm of platform size (ln-platform size) and EPA, and common logarithm of mass (log-mass) upon common logarithm of platform size (log-platform size) and EPA produced results (both r2=.70) and estimates of original size similar to one another. However, for 10 of 25 Nobles Pond specimens the natural-logarithm estimate and for 11 of 25 specimens the common-logarithm estimate of original mass was less than observed mass of the used and discarded specimen. In several other cases, estimates were about 40 g, a figure that considerably exceeds what we, albeit subjectively, consider reasonable estimates of original mass and much exceeds the mean original mass of 19.1 g in Bohush’s experimental assemblage designed to replicate Nobles Pond blanks. That is, a considerable number of sample estimates of original mass from Clarkson and Hiscock data were impossibly low at face value and others were questionably high (for both sets of estimates, coefficient of variation=96.7).

Better results were obtained by confining analysis of Clarkson and Hiscock’s dataset to flakes that did not exceed the maximum Nobles Pond platform area of approximately 81 mm2. Regressing either ln-mass or log-mass upon similarly scaled platform area and EPA produced results of lower explanatory value (both r2=.54). Even then, five of 25 and six of 25 Nobles Pond estimates, for natural-logarithm and common-logarithm variables respectively, were less than observed mass, and some other estimates of original mass exceeded 100 g (for both sets of estimates, coefficient of variation slightly exceeded 121).

We conclude that Clarkson and Hiscock’s dataset, useful for other purposes, is a poor model for estimating original mass of Nobles Pond specimens; not all experimental data are equally good models for all empirical assemblages. Despite measuring platform area precisely using 3D scan data as in our Nobles Pond dataset, Clarkson and Hiscock’s results resemble earlier research on platform allometry. For instance, Davis and Shea (1998:607), Shott et al. (2000:880) and Wilson and Andrefsky (2008:91) reported statistically robust correlations between platform area and mass or other measures of flake size but also estimates for individual specimens that in some cases differed substantially from empirical values. In recent controlled experiments flake mass increased at a higher rate with measures of platform size as EPA increased (Dibble and Rezek 2009:1950, Fig. 6), suggesting nonlinear interaction effects between platform size and EPA. In scanning of experimental flakes, Morales, Lorenzo, and Vergès (2015:12 and Fig. 1) reported unacceptable size and reduction estimates in 20% of cases and suspected high measurement error in EPA. Differences in toolstone, and core size and form also may influence results. There are sources of variation in flake size poorly controlled by platform-allometry estimates that sometimes produce considerable, unpredictable error and counsel further research in this area before using such methods.

Accordingly, we declined to estimate original mass from precisely measured platform area, instead using mass-estimation methods described in the article.

On Multifactorial Estimation

Uncertainty arising from different measures of original size and degree of reduction of stone tools forces archaeologists either to make ad hoc choices among measures or to use the information that inheres in all of them to derive an overall reduction measure. Physical anthropologists confronted a similar dilemma when different skeletal-aging methods (e.g., cranial suture fusion, pubic metamorphosis) yielded significantly different estimates of age-at-death. Lovejoy et al. reasoned that if “more than one criterion is available from which to assess skeletal age at death, then clearly all should be employed” (1985:3), and submitted several independent estimates to principal-components analysis (PCA) of correlation matrices, the resulting first factor or principal component (PC) “assumed to represent” (1985:4) the dimension of interest, true chronological age in their case. Each measure was multiplied by its correlation with PC1, and the sum of products averaged as “multifactorial age.” The method was supported in a controlled skeletal sample where age at death was known independently.

Paleodemographers cricitized Lovejoy et al.’s multifactorial method on several grounds. It involves a biological version of seriation that orders individuals “in a sequence of increasing age” (Lovejoy et al. 1985:3) or increasing wear or attrition in skeletal age estimators. This “seriation” occurs before placing individuals in ranges of inferred age, so might compromise independence between age estimators (Taylor 2013:76) and engages the problem of “age-mimicry,” in which reference samples influence age estimation (Konigsberg and Frankenberg 2002:304).

However, the relevance of these criticisms is unclear in archaeological context, for two reasons. First, forensic age estimation is based on what physical anthropologists call “transition analysis,” the search for natural or empirical boundaries between adjacent age intervals, not the simple continuous variables used here. Second, critics’ alternatives to the multifactorial method (e.g. maximum-likelihood estimation [Konigsberg and Frankenberg 2002] and Bayesian multifactorial estimation [Uhl 2013]) reduce inherently continuous variables to discrete values in ordinal series (Taylor 2013:78). Using them would lower measurement scales, at odds with the continuous nature of both reduction and curation.