American Antiquity (2026), 1–4
doi:10.1017/aaq.2026.10176

SAA
Society for American Archaeology

COMMENT

Comment: Correcting a New Method for Classifying Dart and Arrow Projectile Points

Randall Haas and Robert L. Kelly

Department of Anthropology, University of Wyoming, Laramie, WY, USA
Corresponding author: Randall Haas; Email: whaas@uwyo.edu

(Received 1 December 2025; revised 14 January 2026; accepted 19 January 2026)

Buchanan and colleagues (2026) recently present a new method for classifying dart and arrow projectile points. Following Thomas (1978), the authors use metric attributes from known arrow and dart points in museum collections. This new method deploys classic logistic regression and new statistical techniques, including machine learning and additive modeling within a Bayesian framework. Perhaps their most striking empirical finding is the reclassification of Elko (3.5–2.0 cal ka), Gatecliff Split-Stem (5.8–3.5 cal ka), and Northern Side-Notched (9.0–5.8 cal ka) points (Thomas 2013) as arrow points (Buchanan et al. 2026:Table 1). If valid, this finding would extend the origins of North American archery technology by at least 7,000 years to 9 cal ka, substantially altering several social-evolution models for the Americas (Bettinger 2015; Flores-Blanco et al. 2024).

To independently evaluate Buchanan and coworkers' (2026) finding, we cull direct radiocarbon dates on known archery and atlatl artifacts in North America from the Canadian Archaeological Radiocarbon Database (Kelly et al. 2022). Excluding dates with error terms of ≥200 years, the data include 106 dates—75 atlatl dates and 31 archery dates—ranging between 13 and 0 cal ka. Calibrated dates (Haslett and Parnell 2008; Reimer et al. 2020) for atlatl artifacts produce a 95% range of 12.9–0.7 cal ka (Figure 1). In contrast, archery artifacts are more recent and constrained to a 95% range of 1.4–0 cal ka. The lower bound of archery dates thus falls 7,000 years shy of the lower bound anticipated by Buchanan and colleagues' model, causing concern for the validity of their classification.

To evaluate their method, we first visually inspect the data (Figure 2). Mean width and thickness values of the three point types align most closely with atlatl-point metrics (Figures 2a–b). In contrast, mean length values of the Great Basin point types align most closely with arrowhead lengths. However, resharpening effects confound the utility of length measurements for classification (Shott 1997:94). We furthermore note that hafted points, such as those that form the control sample in Buchanan and colleagues' analysis, are likely to retain more utility—that is, usable material—than the points archaeologists commonly recover, which are more likely to have been exhausted and discarded. The inclusion of point length thus biases their classification toward arrowhead estimation. Although Buchanan and colleagues engage this concern, their treatment does not address the fundamental issue.

Next, we consider the logistic regression method. Logistic regression takes the form,

$$p(x) = \frac{1}{1 + e^{-(x-\mu)/s}},$$

where μ is the location parameter, s is the scale parameter, e is Euler's number (2.718), and x is a predictor variable—in our case, some measure or measures of projectile-point size. Because width and thickness are strongly correlated in the control sample ($t = 11.9$; $p < 0.01$) and because thickness values are missing for some artifacts, we focus on width as the predictor variable in our regression (see also

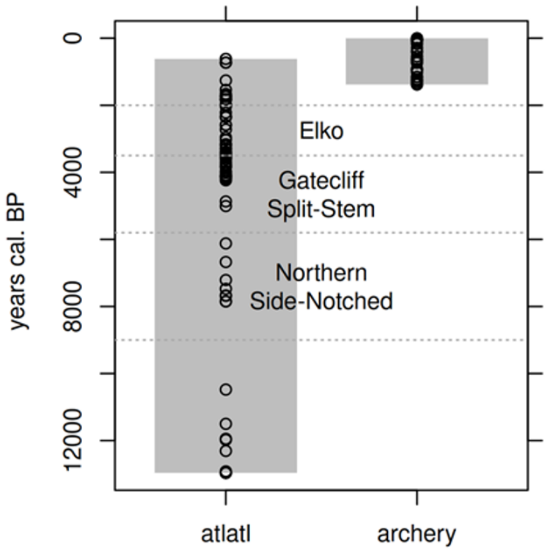

Figure 1. Radiocarbon-dated atlatl ($n = 75$) and archery ($n = 31$) artifacts in North America with temporal comparison to three Great Basin projectile-point types identified by Buchanan and colleagues (2026) as arrow points. The radiocarbon data from known artifact types conflict with their identifications. Dots indicate most-likely calibrated dates of individual specimens. Gray boxes define date ranges for the two technologies.

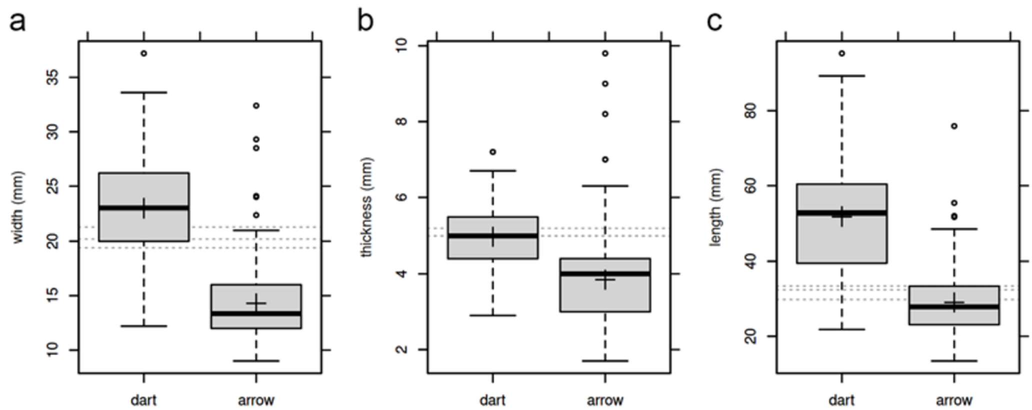

Figure 2. Comparison of control and unknown projectile point metrics, including (a) width, (b) thickness, and (c) length measurements. Boxplots show the distribution of values for control atlatl dart and archery arrow points. + indicates mean values. Dotted gray lines show the mean values for the specimens in question including Elko, Gatecliff Split-Stem, and Northern Side-Notched points of the Great Basin. The widths and thicknesses of these types are consistent with dart points, and the lengths are consistent with arrow points.

Thomas 1978:470). Width values are log transformed for consistency with the assumption of Gaussian error or normality. Buchanan and colleagues do not take such steps to account for non-normality or collinearity, which are known to compromise parameter estimation in the additive modeling and machine-learning approaches that they deploy (Dormann et al. 2012).

To account for imbalance in the frequencies of known dart and arrow points, we down-sample arrowhead width measurements ($n = 220$) to the frequency of dart-point measurements ($n = 53$). To evaluate sample variance, we repeat the sampling procedure 1,000 times. All calculations are performed in the R statistical computing environment (R Core Team 2024).

Our logistic regressions reveal a consistently strong relationship between projectile technology and projectile-point width ($z{\sim}5.5$; $p < 0.01$). Across all iterations, arrow/dart-point width thresholds, μ,

https://doi.org/10.1017/aaq.2026.10176 Published online by Cambridge University Press

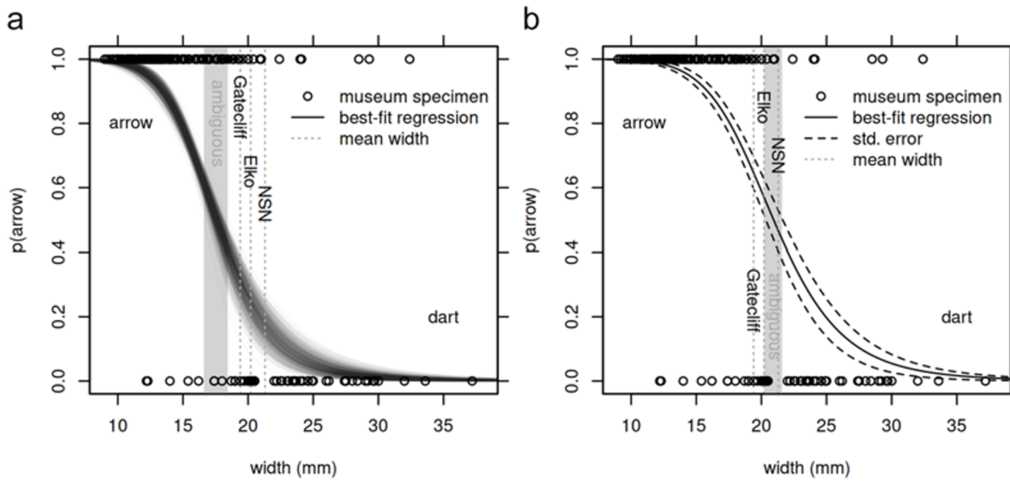

Figure 3. Logistic regression of projectile system as a function of projectile-point width: (a) 1,000 iterations balancing arrow and dart sample sizes; (b) regression using raw, unbalanced data. Gray areas, labeled "ambiguous," show the threshold widths outside of which artifacts are more likely arrow or dart points. Dotted gray lines show the mean width values for Elko, Gatecliff Split-Stem, and Northern Side-Notched points. The model using balanced data identifies the width threshold, μ, at 16.6–18.5 mm and the three Great Basin point types as dart points. The model using raw, unbalanced data spuriously identifies the width threshold, μ, at 20.2–21.6 mm and the three Great Basin point types as arrows or ambiguous points.

range between 16.6 and 18.5 mm. The model furthermore estimates the following arrow probabilities based on the mean widths presented in Buchanan and colleagues (2026:Table 1): Elko 10%–38%, Gatecliff 14%–43%, and Northern Side-Notched 5%–31% (Figure 3a). In other words, the model consistently identifies these point forms as more likely to have been dart points than arrow points [*p(arrow)* < 0.5].

In contrast, running the regression on the full, unbalanced control dataset inflates the logistic curve to favor arrow estimates, placing the arrow/dart-point width threshold between 20.2 and 20.6 mm (Figure 3b). Thus, the three Great Basin point types in question are spuriously identified as arrow points or are ambiguous, aligning more closely with Buchanan et alia's findings. This leads us to suspect that their balancing method is inadequate and directionally biases parameter estimation. The fundamental issue appears to lie in the assumption that the proportion of dart and arrow points in the control sample (19:81 dart:arrow) reflects the proportion points found in field investigations. However, the highly skewed proportion in the model is unlikely to reflect the reality of archaeologically recovered projectile points, for which a vague prior of 50:50 would be more appropriate, lacking additional information.

Finally, Buchanan and colleagues deploy a spline function in their regression, which necessarily overfits a non-monotonic model to a monotonic phenomenon. There is no theoretical reason to suppose that point size non-monotonically predicts projectile system. This error invalidates their cross-validation analysis and precludes formal comparison with the corrected model presented here.

Analysts looking for a simple but accurate method to classify projectile points would be better served by assuming a threshold width of 16.6–18.5 mm, such that narrower points are more likely arrow points, wider points are more likely dart points, and points within the range are ambiguous. For probabilistic estimation, one could use a standard calculator or spreadsheet to enter projectile-point widths into the following logistic equation, which incorporates the median μ and s parameter estimates derived from the current analysis:

$$p\,(arrow) = \frac{1}{1 + e^{-(log(w) - log(17.6))/-0.122}},$$

https://doi.org/10.1017/aaq.2026.10176 Published online by Cambridge University Press

where *p(arrow)* is the probability that a given point is an arrowhead, *e* is Euler's number (2.718), and *w* is the measured width in mm of the projectile point. Here, *p(dart)* is simply 1 − *p(arrow)*. For example, entering a width of 20 mm returns a *p(arrow)* of 26%, thus indicating a 74% chance that the projectile point is a dart point. Additional uncertainty could be considered by exploring *p(arrow)* across the range of μ and *s* estimates of 16.6–18.5 mm and −0.174 to −0.087, respectively.

Although we commend Buchanan and colleagues' (2026) use of logistic regression for probabilistic projectile-point classification, we caution against wholesale application of their method. The use of projectile-point length, unbalanced control data, and collinear predictors all inflate parameter estimation, directionally biasing estimation in favor of arrows. The use of non-normalized data and spline functions further compromise the method. More substantively, Buchanan and colleagues' findings do not give cause to redefine the history of projectile technology and its consequences in the Americas.

References Cited

Bettinger, Robert L. 2015. *Orderly Anarchy: Sociopolitical Evolution in Aboriginal California*. University of California Press, Oakland.

Buchanan, Briggs, Marcus J. Hamilton, and Robert S. Walker. 2026. A New Method for Classifying Dart and Arrow Projectile Points. *American Antiquity* 91(1):168–182. https://doi.org/10.1017/aaq.2025.10109.

Dormann, Carsten F., Jane Elith, Sven Bacher, Carsten Buchmann, Gudrun Carl, Gabriel Carré, Jaime R. García Marquéz, et al. 2012. Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance. *Ecography* 36(1):27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x.

Flores-Blanco, Luis, Lucero Cuellar, Mark Aldenderfer, Charles Stanish, and Randall Haas. 2024. Did Archery Technology Precipitate Complexity in the Titicaca Basin? A Metric Analysis of Projectile Points, 11–1 ka. *Quaternary International* 704:17–33. https://doi.org/10.1016/j.quaint.2023.10.012.

Haslett, John, and Andrew C. Parnell. 2008. A Simple Monotone Process with Application to Radiocarbon-Dated Depth Chronologies. *Journal of the Royal Statistical Society Series C: Applied Statistics* 57(4):399–418. https://doi.org/10.1111/j.1467-9876.2008.00623.x.

Kelly, Robert L., Madeline E. Mackie, Erick Robinson, Jack Meyer, Michael Berry, Matthew Boulanger, Brian F. Codding, et al. 2022. A New Radiocarbon Database for the Lower 48 States. *American Antiquity* 87(3):581–590. https://doi.org/10.1017/aaq.2021.157.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/, accessed January 14, 2026.

Reimer, Paula J., William E. N. Austin, Edouard Bard, Alex Bayliss, Paul G. Blackwell, Christopher Bronk Ramsey, Martin Butzin, et al. 2020. The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon* 62(4):725–757. https://doi.org/10.1017/rdc.2020.41.

Shott, Michael J. 1997. Stones and Shafts Redux: The Metric Discrimination of Chipped-Stone Dart and Arrow Points. *American Antiquity* 62(1):86–101. https://doi.org/10.2307/282380.

Thomas, David Hurst. 1978. Arrowheads and Atlatl Darts: How the Stones Got the Shaft. *American Antiquity* 43(3):461–472. https://doi.org/10.2307/279405.

Thomas, David Hurst. 2013. Great Basin Projectile Point Typology: Still Relevant? *Journal of California and Great Basin Anthropology* 33(2):133–152.

Cite this article: Haas, Randall, and Robert L. Kelly. 2026. Comment: Correcting a New Method for Classifying Dart and Arrow Projectile Points. *American Antiquity.* https://doi.org/10.1017/aaq.2026.10176.