The Virtual Goniometer: Demonstrating a new method for measuring angles on archaeological materials using fragmentary bone [*]

Katrina Yezzi-Woodley[†]   Jeff Calder[‡]   Peter J. Olver[‡]   Paige Cody [§]

Thomas Huffstutler [‡]   Alexander Terwilliger [‡]   J. Anne Melton[§]

Martha Tappen[§]   Reed Coil[¶]   Gilbert Tostevin [§]

April 16, 2021

Abstract

The contact goniometer is a commonly used tool in archaeological analysis, despite suffering from a number of shortcomings due to the physical interaction between the measuring implement, the object being measured, and the individual taking the measurements. However, lacking a simple and efficient alternative, researchers in a variety of fields continue to use the contact goniometer to this day. In this paper, we present a new goniometric method that we call the *virtual goniometer*, which takes angle measurements on a 3D model of an object. The virtual goniometer allows for rapid data collection, and for the measurement of many angles that cannot be physically accessed by a manual goniometer. Using fracture angle measurements on bone fragments, we compare the intra-observer variability of the manual and virtual goniometers, and find that the virtual goniometer is far more consistent and reliable. Furthermore, the virtual goniometer allows for precise replication of angle measurements, even among multiple users, which is important for reproducibility of goniometric-based research. The virtual goniometer is available as a plug-in in the open source mesh processing packages Meshlab and Blender, making it easily accessible to researchers exploring the potential for goniometry to improve archaeological methods and address anthropological questions.

Keywords— goniometer, taphonomy, zooarchaeology, fracture angle, bone fragments, faunal, lithics

1   Introduction

Goniometry is an important aspect of archaeological and zooarchaeological analysis. The primary tool for studying angles on objects, such as bone fragments or lithics, is the pocket, contact goniometer, in essence

---

a metal protractor with a rotating arm. The contact goniometer was originally designed in the 1780s to measure the angles on crystals (Burchard, 1998) and first appeared in archaeology when Barnes (1939) used the instrument to differentiate anthropogenically produced stone tools from naturally-occurring conchoidally fractured rocks.

Lithic analysts have been measuring angles on stone tools since Barnes (1939) to address a number of questions including but not limited to: tool function (e.g. Gould, Koster, & Sontz, 1971; Wilkins et al., 2017; Wilmsen, 1968), retouch intensity (e.g. Kuhn, 1990), and technological and behavioral variability (e.g. Dibble, 1997; Nigst, 2012; Režek, Dibble, McPherron, Braun, & Lin, 2018; Scerri, Gravina, Blinkhorn, & Delagnes, 2016; Tostevin, 2003). Taphonomic applications began when Capaldo and Blumenschine (1994) used the goniometer to distinguish bones broken by carnivores from those broken by hominins. They extrapolated directly from lithic methods using the goniometer to measure the internal platform angle on models of bone flakes created by taking impressions of notches and associated flake scars on experimentally broken bone. In 2006, Alcántara-García et al. (2006) introduced a method for identifying actors of breakage by using the goniometer to measure fracture angles — meaning the angle of transition from the periosteal surface to the fracture surface — on long bone shaft fragments. Prior to this, fracture angles were assessed by eye and categorized as oblique, right, or both, as a means of distinguishing green breaks from dry breaks (Villa & Mahieu, 1991).

Using a contact goniometer in the analysis of fragmentary faunal assemblages is gaining traction because it permits researchers to collect seemingly more reliable, quantitative data, and opens the possibility for other avenues of analysis to address questions related to hominin and carnivore interactions at important paleoanthropological sites (Coil, Tappen, & Yezzi-Woodley, 2017; De Juana & Dominguez-Rodrigo, 2011; Domínguez-Rodrigo & Barba, 2006; Moclán, Domínguez-Rodrigo, & Yravedra, 2019; Moclán et al., 2020; Pickering, Domínguez-Rodrigo, Egeland, & Brain, 2005; Pickering & Egeland, 2006). However, the reliability of the contact goniometer and the full range of possibilities for goniometry in faunal analysis has not been fully explored. Furthermore, the principles of goniometry are the same whether applied to bone fragments or stone tools and, within lithics, the contact goniometer's reliability has come under question and alternative methods have been proposed (Dibble & Bernard, 1980; Morales, Lorenzo, & Vergès, 2015). The limitations imposed on the contact goniometer due to its physical interaction with the target object prevents a more rigorous exploration of the potential for goniometry in archaeological inquiries. A more flexible and precise tool is required.

Here we introduce a new method for taking angle measurements, called the *virtual goniometer*. Though this new method assumes that 3D models are readily available for use and does not account for the time and expense of acquiring the models, scanning methods are rapidly becoming more efficient and cost-effective (Adamopoulos, Rinaudo, & Ardissono, 2021; Das, Murmann, Cohrn, & Raskar, 2017; Maté-González, Aramendi, González-Aguilera, & Yravedra, 2017; Porter, Huber, Hoyer, & Floss, 2016; Porter, Roussel, & Soressi, 2016; Sapirstein & Murray, 2017). Being able to scan a larger number of samples allows researchers within archaeology to expand the possibilities for research by developing digital tools that can extract data from 3D models, such as the virtual goniometer (Archer, Gunz, van Niekerk, Henshilwood, & McPherron, 2015; Archer, Pop, Gunz, & McPherron, 2016; Valletta, Smilansky, Goring-Morris, & Grosman, 2020).

The virtual goniometer can effectively measure angles across the range of values (acute to obtuse) which is not the case for extant physical tools. For example, the steep exterior platform angles (EPA) on lithics are better measured with a contact goniometer than with the caliper-goniometer (Dibble & Bernard, 1980) because the steep apex angle makes the set distance between the points of the caliper and the bar too short for consistent measurement whereas the arms of the contact goniometer can capture more of the dorsal surface for a meaningful EPA as it would be visible to the knapper. On the other hand, very acute angles of cutting edges cannot be measured with contact goniometers and are better measured with the caliper-goniometer (Key & Lycett, 2015). The virtual goniometer can cover both ranges of edge values at the same time, and

that is within one artifact class.

Due to how the virtual goniometer measures angles, it can be used to measure 3D models of any object and can be applied across archaeological contexts and materials, thus extending the use of 3D models beyond digital preservation and creating opportunities for new avenues of research within anthropology . The virtual goniometer is an enabling technology that allows researchers to control how measurements are taken which is key to creating protocols that extract anthropologically useful information and can be easily replicated in independent studies. With these future studies in mind, we have implemented the virtual goniometer as a plug-in in the open source mesh processing software packages Meshlab and Blender, making the method widely available to researchers in the field.

2 Materials and Methods

We compared the manual and virtual goniometers for computing fracture angles on a sample of bone fragments consisting of 537 breaks taken from 86 appendicular long bone shaft fragments (≥ 2 cm maximum dimension) randomly chosen from a collection of experimentally broken *Cervus canadensis nelsoni* (Rocky Mountain elk) and *Odocoileus virginianus* (white-tailed deer) limb bones. Fragments were scanned using a medical CT scanner at the University of Minnesota's Center for Magnetic Resonance Research (CMRR) (slice thickness: 0.6, reconstruction interval: 0.6 mm, KV: 80, MA:28, rotation time: .05 sec, pitch: 0.8, algorithm: bone window, convolution kernel: B60f sharp) and then surfaced using MATLAB. The break surfaces on each fragment were manually subdivided into separate break planes (Gifford-Gonzalez, 1989; Haynes, 1983; Pickering et al., 2005). All breaks were measured, and we did not impose a minimum break length requirement. We measured the fracture angle — which is defined as the angle of transition between the periosteal surface and break surface — on each break face of each fragment (Alcántara-García et al., 2006; Capaldo & Blumenschine, 1994; Villa & Mahieu, 1991). Following the method established by Alcántara-García et al. (2006) and further described by Pickering et al. (2005), we chose to measure at the center along the break length.

Each break was measured using both a contact goniometer ("man") and the virtual goniometer using the Meshlab plug-in. Two methods were employed using the virtual goniometer. We refer to these methods as the *click-and-drag method* ("drag") and the *xyz method* ("xyz"). We provide a brief overview of the two virtual goniometer methods here. A more detailed description can be found in section 2.1. The fragments were first measured using the drag method, whereby the user inputs the 3D model into Meshlab or Blender (in this case Meshlab) and, using the mouse, clicks the location on the fracture edge where the measurement is to be taken. The user drags the mouse to select a region of the surface mesh surrounding the selected location, which we call a *patch*, to be used in the angle calculation (see Fig. 1). After the first round of drag angle measurements, denoted θ_{drag}, screen shots were taken of all the colorized models to create a 2D map of the measurements (see Fig. 2). This served as a visual guide for subsequent drag angle measurements, which we denote by φ_{drag} and ψ_{drag} and the three manual angle measurements, $\theta_{man}, \varphi_{man}, \psi_{man}$, taken using the contact goniometer.

For each measurement taken using the virtual goniometer, data are output into an automatically generated .csv file, including the radius of the circular patch used for the measurement and the xyz-coordinates of the center of the patch. These data can be input directly into Meshlab (or Blender) as an alternative to interfacing directly with the 3D mesh using the mouse. (See 2.1 for a detailed description.) For the xyz method we replicated the first set of drag measurements. Thus, the xyz method was only executed twice, so $\theta_{xyz} = \theta_{drag}$. The values for the radius and location (i.e. xyz-coordinates) of θ_{drag} were input into the plug-in for the xyz method to produce two further angle measurements $\varphi_{xyz}, \psi_{xyz}$).

In total, the same person (KYW) measured the angle of each break eight times (3 manual, 3 drag, and 2 xyz) thereby allowing us to test for intra-observer error. Following the results of the intra-observer test,

3

it was deemed not necessary to run inter-observer tests for reasons which will be discussed in the results section. We tested for accuracy and precision in all methods, and demonstrate the ease of replicability when using the virtual goniometer. Accuracy is the difference between the true angle and the measured angle. Precision is the consistency of repeated measurements. Using additional data automatically provided by the two virtual methods, we analyzed the degree to which the location of the measurement, the distance from the edge, and the number of mesh points used impact how the angle measurement varies.

2.1 How the virtual goniometer works

The first step in the process is to scan the object of interest to produce a meshed surface. In most of our work, the mesh consists of a large number of triangles that approximate the bounding surface of the solid object. The common vertices of the triangles will be referred to as *mesh points*. Alternatively, the scanned surface could just be represented by an unordered "point cloud" consisting of mesh points without specification of the associated triangles. Typically the number of mesh points (triangles) on our scans ranges from 20,250 to 288,029.

In order to implement our virtual goniometer angle measurements, in addition to the mesh points representing the surface, we also require a unit outward *normal* at each mesh point, meaning a vector of length one that points exterior to the object and is normal — meaning perpendicular — to the tangent plane of the surface at its position. If the surface is a triangulated mesh, one can readily compute the normal vector to each triangle, and then average the normal vectors over nearby triangles to determine the normal at a given mesh point. In the case of a point cloud, one can employ local Principal Component Analysis (PCA) or other convenient methods (Bishop, 2006) to compute the normal. With the mesh points and their normals in hand, we are ready to apply the virtual goniometer to compute angles at selected points on the surface.

The first ingredient of each angle measurement is the specification[1] of one mesh point on the surface. The location is assumed to be on, or at least close to, a break edge or interface where an angle measurement has meaning. Choosing a location far away from an edge can lead to spurious angle measurements that are of no use to the studies for which the virtual goniometer is designed. In our implementation, the mesh point specification can be done in two ways: in the first, which we call the *click-and-drag method*, the user clicks on the displayed surface mesh in Meshlab to choose the point; alternatively, in the *xyz method*, the user enters the xyz-coordinates directly. The second method is particularly useful when reproducing or re-

---

[1]We are actively developing methods for automatically determining the break edges on objects of interest, following which one can completely automate the corresponding angle measurements along the edges.
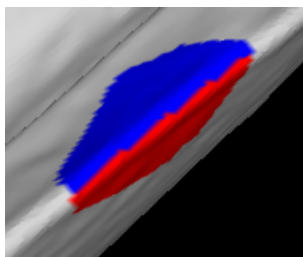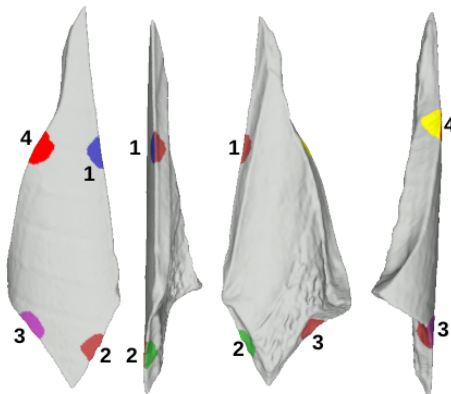

Figure 1: Angle measurement

Figure 2: Example of the 2D map

evaluating measurements since it avoids any inherent uncertainty in users clicking on different points each time.

The second ingredient is a small section of the surface, referred to as a *patch*, that is more or less centered on the chosen mesh point. Consequently, one needs to specify the size or *radius $r > 0$* of the patch. When using the click-and-drag method, the user clicks the desired location, and then drags the cursor to select the radius to be used. While dragging, the software interactively colors the patch to enable the user to see how large or small it will be, and once the cursor is released, the patch is chosen for the angle measurement. Alternatively, the user can input a specified radius. In more detail, once the radius $r > 0$ has been specified, the patch consists of all surface mesh points that are within a distance *r* of the specified point. In our Meshlab implementation, the distance is measured along the surface itself, known as *geodesic distance*. In Blender, distance is simply the standard Euclidean distance in the ambient three-dimensional space. The differences between the two distance measurements only slightly affect the specification of the patch, and produce very little difference between the resulting angle measurements.

The heart of our goniometer algorithm is to separate or *cluster* the mesh points in the patch into two subsets. (The complete mathematical details can be found in the next Section 2.2; here we give a simplified description of the underlying ideas.) Each of these subsets should contain the mesh points that lie on one of the sides of the break edge that is assumed to be passing through (or near) the specified center point of the patch. The clustering algorithm we employ primarily relies on the surface normals at each mesh point in the patch. For example, suppose the surface has the form of a crystal and consists of two flat planes intersecting along a line, which is the edge. Then all the points on one side of the edge would have a common normal direction, while all the points on the other side would have a different common normal direction, and hence the normal directions serve to cluster the points lying on the two sides of the edge. More generally, one seeks to cluster the normals for the patch points into two classes, based on how close they are to each other. The mesh points in the patch whose normal is in the first class are then deemed to lie on one side of the break edge, while those in the second class lie on the other side. Our algorithm is based on the random one-dimensional projection clustering methods introduced in Han and Boutin (2014); Yellamraju and Boutin (2018), and consists of projecting the normals in a certain intrinsic direction — the *binormal direction* that is tangent to the surface and perpendicular to the edge — and then splitting the resulting one-dimensional projected normal data into the two classes, which can be easily and quickly done. This relatively simple algorithm is fast, works very well in practice, and outperforms or is close to the performance level of much more sophisticated clustering algorithms.

However, the clustering based entirely on normals described in the previous paragraph can be significantly improved by taking into account the distances between mesh points, keeping in mind that points lying on one side of a break edge should mostly be closer to each other than to those on the opposite side. Thus, the full clustering algorithm uses both the normal data and the interpoint distance data. There is a *tuning parameter λ* that weighs the relative importance of these two data components, and the user can, if desired, alter its value if the initial clustering and resulting angle measurement is suboptimal.

Once the clustering algorithm is completed — which happens almost instantaneously — the selected patch is color-coded into two contrasting colors showing the individual clusters representing the two sides of the break curve. See Figure 5 below for a sketch of such a bicolored surface patch. The colors rotate through a pallet list of pairs of contrasting colors; see Figure 3 for examples of the resulting visual output in Meshlab.

The algorithm finally computes the break angle by approximating each of the two mesh point clusters by a two-dimensional plane, using standard methods based on Principal Component Analysis (PCA). Thus, on either side of the break edge, the object's surface has been locally approximated by a plane. The angle between these two planes, computed by a standard trigonometric formula, is deemed to be the Virtual Goniometer angle measurement at the center of the patch.

Further measurements can be taken at or near the original location; see Figure 3b. The user can also choose to advance to a new location at which point the colors change; see Figures 3b and 3c. Numerical data for each measurement are automatically recorded in a `.csv` file. The plug-in documentation which provides detailed, step-by-step instructions is available at `https://amaaze.umn.edu/`
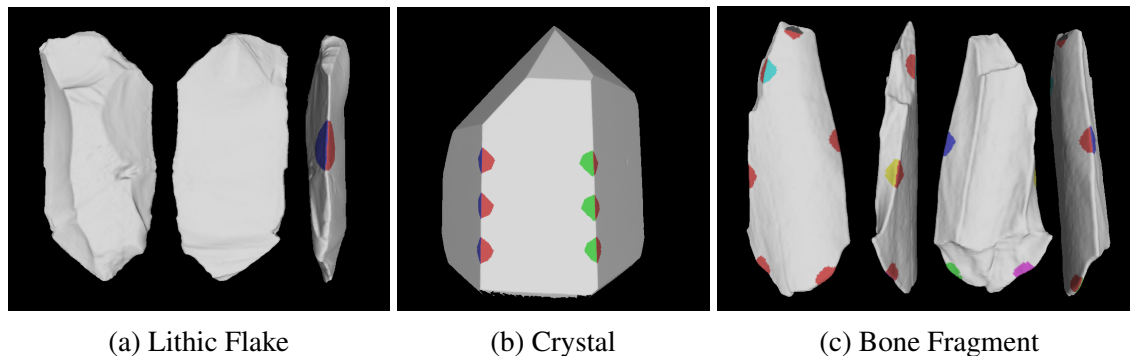

(a) Lithic Flake   (b) Crystal   (c) Bone Fragment

Figure 3: Examples of the Virtual Goniometer on Different Materials

2.2   The math behind the virtual goniometer

In this section we describe the mathematical algorithms used to design the virtual goniometer. This section is for those interested in the mathematical details that make it possible to fully reproduce our work in independent research.

To describe the algorithm mathematically, we first note that all vectors are considered as column vectors. There are two key matrices that serve as the input to the algorithm: $\mathbf{X}$ will represent all the mesh points in the patch, while $\mathbf{N}$ will represent their corresponding outward unit normal vectors, while λ is the aforementioned tuning parameter. The Virtual Goniometer is summarized in Algorithm 1, whose output is the angle measurement θ in degrees, and a goodness of fit $\varepsilon > 0$.

More precisely, let n be the number of mesh points in the patch. We use $\mathbb{1} = [1, \ldots, 1]^T$ to denote the column vector with n entries all equal to one, and $\|\mathbf{x}\|$ to denote the Euclidean norm of the vector $\mathbf{x}$. Let $\mathbf{X}$ be the $3 \times n$ matrix whose i^{th} column is the vector of x, y, z coordinates of the ith mesh point in the selected patch. Let $\mathbf{N}$ be the $3 \times n$ matrix whose i^{th} column is the outward unit normal vector to the surface at the ith mesh point.

As noted in section 2.1, The Virtual Goniometer Algorithm uses both the normal vectors $\mathbf{N}$ and the mesh points $\mathbf{X}$ to segment the patch into two regions. The tuning parameter λ controls the tradeoff between how much to rely on the normals versus the mesh points. Setting $\lambda = 0$ results in clustering using only the normal vectors $\mathbf{N}$. If the surface is noisy, this can give a poor segmentation, since some normal vectors could point in a similar direction even if they are on opposite sides of a break. Increasing λ encourages the segmentation to put points that are close together into the same region, and can help to improve the segmentation on noisy meshes. Figure 4 shows the effect of λ on the segmentation, and there are generally three instances, when it needs to be adjusted: (4a) sharp curves in the ridge; (4b) subtle ridges, usually associated with obtuse angles; and (4c) rugose surfaces. In our implementation, we take $\lambda = 2$ as the default value of the tuning parameter and find that we only need to change this parameter for a small number of measurements. Moreover, the user can easily learn when and how to adjust it.

Let us describe the individual steps in our algorithm in some detail. In Step 3, one can use for $\bar{\mathbf{x}}$ any reasonable notion of centroid of the patch $\mathbf{X}$, and either the mean value of the coordinates or a geodesic centroid work well. In our implementation in Meshlab, we set $\bar{\mathbf{x}}$ to be the mesh point selected by the user in

Algorithm 1 Virtual Goniometer

1: **Input:** Points $\mathbf{X}$, normals $\mathbf{N}$, and tuning parameter $\lambda \geq 0$
2: **Output:** Angle θ and goodness of fit ε
3: $\bar{\mathbf{x}} \leftarrow$ Centroid of points $\mathbf{X}$. {See notes.}
4: $r \leftarrow$ Radius of patch $\mathbf{X}$. {See notes.}
5: $\mathbf{t} \leftarrow$ Eigenvector of $\mathbf{N}\mathbf{N}^T$ with smallest eigenvalue. {Tangent to break curve.}
6: $\mathbf{n} \leftarrow \dfrac{1}{n}\mathbf{N}\mathbb{1}$ {Average outward normal vector.}
7: $\mathbf{b} \leftarrow \dfrac{\mathbf{t} \times \mathbf{n}}{\|\mathbf{t} \times \mathbf{n}\|}$ {Cross product, normalized — binormal vector.}
8: $\mathbf{p} \leftarrow \mathbf{b}^T\left[\mathbf{N}+(\lambda/r)(\mathbf{X}-\bar{\mathbf{x}})\right]$ {Dot product of all rows with $\mathbf{b}$.}
9: $s \leftarrow \text{WithinSS}(\mathbf{p})$ {Find best clustering of 1D data $\mathbf{p}$; see notes.}
10: $\mathbf{X}_- = \mathbf{X}[\mathbf{p} < s]$, $\quad \mathbf{X}_+ = \mathbf{X}[\mathbf{p} \geq s]$ {Split into subpatches divided by break curve.}
11: $(\mathbf{v}_\pm, \varepsilon_\pm) \leftarrow \text{PCA}(\mathbf{X}_\pm)$ {Principal component with smallest variance, mean squared error.}
12: $\theta \leftarrow 180 - \arccos\left(\text{sign}(\mathbf{n}^T\mathbf{v}_+)\,\text{sign}(\mathbf{n}^T\mathbf{v}_-)\,\mathbf{v}_+^T\mathbf{v}_-\right)$ {VG angle — arccos returns degrees.}
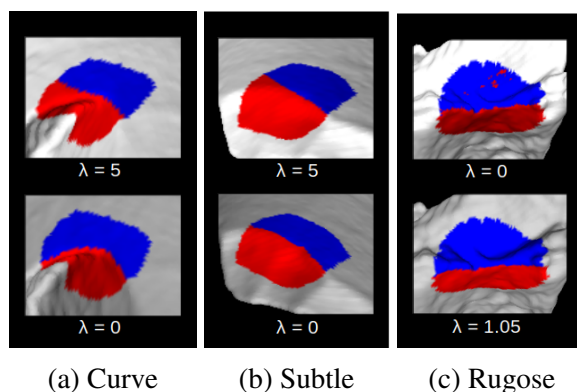13: $\varepsilon \leftarrow \varepsilon_+ + \varepsilon_-$ {Goodness of fit.}

$\lambda = 5$ $\quad\quad$ $\lambda = 5$ $\quad\quad$ $\lambda = 0$

$\lambda = 0$ $\quad\quad$ $\lambda = 0$ $\quad\quad$ $\lambda = 1.05$

(a) Curve $\quad\quad$ (b) Subtle $\quad\quad$ (c) Rugose

Figure 4: Examples of how the tuning parameter λ affects how the measurement is taken.

the click-and-drag selection method or the point specified in the xyz method. In Step 4, we use the geodesic radius of the patch for r; that is the largest geodesic distance from $\bar{\mathbf{x}}$ to any point in $\mathbf{X}$ as measured along the surface. It is also possible to use the Euclidean radius of the patch, as we do in the Blender implementation, and the only difference is a minor change in the effect of the tuning parameter λ. However, the algorithm is not overly sensitive to this effect.

The vector $\mathbf{t}$ produced in Step 5 is to be interpreted as the tangent vector to the "break curve" that separates the two approximately planar regions on the surface. The vector $\mathbf{n}$ in Step 6 is the averaged outward normal vector to the patch. Thus, in Step 7, the vector $\mathbf{b}$ given by normalizing the cross product of the break curve tangent and the surface normal is perpendicular to both and hence can be interpreted as the unit *binormal* vector of the patch $\mathbf{X}$, pointing *across* the break curve, i.e., a unit vector that is both tangent to the surface and normal to the break curve. The binormal vector $\mathbf{b}$ is used to quickly obtain the correct segmentation, as described below. See Figure 5 for a depiction of the vectors $\mathbf{t}, \mathbf{b}$, and $\mathbf{n}$ when the surface is (approximately) planar on either side of the (approximately) straight break edge.

Our clustering method used in Step 9 to divide the surface into two classes was inspired by the random projection clustering methods of Han and Boutin (2014); Yellamraju and Boutin (2018), which involve
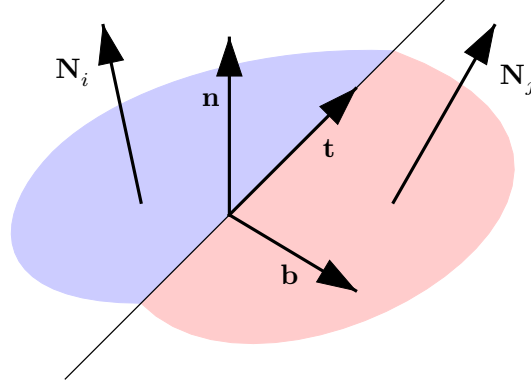
Figure 5: Depiction of a surface patch and the tangent, normal and binormal vectors $\mathbf{t}, \mathbf{n}, \mathbf{b}$ in Algorithm 1. The vector $\mathbf{t}$ is tangent to the break curve, which is a line in this example, $\mathbf{n}$ is the average of the outward normals on both sides of the break edge, denoted $\mathbf{N}_i$ and $\mathbf{N}_j$ in the figure, and the binormal $\mathbf{b}$ is perpendicular to $\mathbf{n}$ and $\mathbf{t}$, and hence points *across* the break edge. The basis for our segmentation algorithm is that the sign of the dot product $\mathbf{b} \cdot \mathbf{N}_i$ will indicate on which side of the break edge a particular meshpoint falls.

repeatedly randomly projecting the data to one dimension, and then using the function `WithinSS` described below to perform the clustering of the resulting one-dimensional projected data. Given one-dimensional data represented by

$$\mathbf{p} = (p_1, p_2, \ldots, p_n) \qquad \text{where} \qquad p_1 \leq p_2 \leq \cdots \leq p_n,$$

this function computes a real number s for which the quantity

$$f(s) = \sum_{p_i \geq s} (p_i - c_1)^2 + \sum_{p_i < s} (p_i - c_2)^2 \qquad \text{where} \qquad c_1 = \frac{\sum_{p_i \geq s} p_i}{\sum_{p_i \geq s} 1}, \qquad c_2 = \frac{\sum_{p_i < s} p_i}{\sum_{p_i < s} 1}.$$

is minimized. The value of s that minimizes $f(s)$ gives the optimal clustering of the one-dimensional data $\mathbf{p} = (p_1, \ldots, p_n)$ into two groups $\{p_i \geq s\}$ and $\{p_i < s\}$. Note that $f(s)$ is constant on each interval $p_i < s \leq p_{i+1}$, and hence we can the global minimizer of f simply by computing $f(p_i)$ for $i = 1, \ldots, n$, and choosing $s = p_i$ that gives the smallest value. (And hence there is also no need to actually sort the data points $\mathbf{p}$.) This highlights the advantage of working with one-dimensional data; it is very simple and fast to compute optimal clusterings.

Step 8 in the Virtual Goniometer Algorithm implements a one-dimensional projection of a λ weighted combination of the unit normals to the mesh points in the patch and the mesh points themselves, shifted by the centroid so as to center the patch at the origin. However, the direction of the projection $\mathbf{b}$ is not random but is very carefully chosen as the *binormal* of the patch. The random projection algorithm advocated in Han and Boutin (2014) also works very well in this application; however, it requires around 100 random projections to obtain reliable results and is thus significantly slower. We have also experimented with other clustering algorithms, such as the hyperspace clustering method of Zhang, Szlam, Wang, and Lerman (2010), which also gives very good results at the expense of longer computation times. Our method presented in Algorithm 1 is very efficient and is suitable for real-time computations in mesh processing software such as Meshlab.

The resulting pair of clusters

$$\mathbf{X}_- = \mathbf{X}[\mathbf{p} < s] = \{\mathbf{X}_i \mid p_i < s\} \qquad \text{and} \qquad X_+\mathbf{X}[\mathbf{p} \geq s] := \{\mathbf{X}_i \mid \mathbf{p}_i \geq s\}$$

contain the mesh points belonging to the two components of the surface patch on either side of the break curve. (However, we do not construct the actual break curve; nor do we require that the centroid or selected location $\bar{\mathbf{x}}$ be thereon.) In Steps 10 and 11, the PCA function returns each principal component, denoted $\mathbf{v}_{\pm}$, with the smallest variance and the corresponding eigenvalues $\varepsilon_{\pm}$, which represents the mean squared error in the fitting. We also experimented with robust versions of PCA (see Lerman and Maunu (2018) for an overview), but did not find the results were any more consistent. Finally, in Step 13, arccos is the inverse cosine function, measured in degrees, not radians, and we use the branch with values between 0 and 180 degrees.

2.3 Other computer-based methods for measuring angles

Recently, in the field of lithic analysis, efforts have shifted in a digital direction (Archer et al., 2016; Grosman, Goldsmith, & Smilansky, 2011; Grosman, Karasik, Harush, & Smilansky, 2014; Grosman, Smikt, & Smilansky, 2008; Valletta et al., 2020; Weiss, 2020; Weiss, Lauer, Wimmer, & Pop, 2018). Archer et al. (2015, 2016) developed an R package that calculates edge angles based on the thickness of the object at a fixed distance from the edge using basic trigonometry. Their method applies PCA to find the principal axes of the lithic, the first determining its long axis, the second its width, defined as the furthest extent of the object in that principal direction, and the third its thickness, defined as the distance between corresponding points on each biface. The angle at an edge point is then calculated using the isosceles triangle in the plane perpendicular to the edge, whose apex is the edge point and whose base equals the thickness at a specified distance from the edge point. As pointed out in the code description by Pop (2019), the function will not work if the plane intersects more than one edge. Therefore, the function depends on the object being of a particular shape, is sensitive to the location of the points where measurements are taken, and cannot be easily applied to other tool types or bone fragments which have break edges that are not similar to those found on lithics, such as spiral breaks. Because the angle calculation relies on only three points, it is highly sensitive to small topographical changes on the object. The virtual goniometer, however, works on completely general digitized solid objects and uses all the mesh points within the patch to define the angle of intersection between the two faces and is therefore not sensitive to small topographical deviations.

Valletta et al. (2020) developed an angle measurement procedure available in a stand-alone software program. This program uses 3D models and calculates a mean value of the angle measurements taken from a number of selected points along an edge (Valletta et al., 2020). They use a cylindrical area that encompasses the entire length of the ridge, or break edge, and average the data along that length to fit two planes on either side of the ridge. This is useful when the edges are straight and uniform. However, many lithic elements and bone fragments do not conform to this ideal.

These two methods were designed specifically for lithic analysis, particularly for blades and bifaces, and are restricted to objects of this general shape. On the other hand, the virtual goniometer can be used to measure angles on completely general digitized objects arising in a very broad range of applications.

3 Results

3.1 Summary statistics for all methods

Of the 537 breaks in our randomly selected sample of bone fragments, 500 (93.1%) could be measured manually using the contact goniometer, while the other 37 breaks could not be physically measured. For 34 of those breaks, one or both arms of the contact goniometer were blocked from contacting the break face or periosteal face. The 3 remaining breaks that could not be measured manually came from a fragment that suffered lab damage after being scanned. The 3D mesh constructed prior to the lab break made it possible

to apply the virtual goniometer to these samples that were unavailable for manual measurements. When comparing the manual measurements to the click-and-drag and xyz methods, we only used the 500 breaks that could be measured by all three methods.

To test variability, we computed three angle measurements for each of the three methods, which we denote by θ, φ, ψ and use subscripts man, drag, xyz to indicate which method is employed. When using the manual and drag method, the location of the measurement is selected by eye. For the xyz method, the user enters the x, y, z coordinates directly.

To assess how much each method varied, we calculated the Intra-Observer Variability (IOV) for the three angle measurements (θ, φ, ψ) for each break under the three different methods. The IOV is the average of the absolute value of the differences between each of the three angle measurements, all taken at the same location:

$$IOV_{man} = \frac{|\theta_{man} - \varphi_{man}| + |\theta_{man} - \psi_{man}| + |\varphi_{man} - \psi_{man}|}{3},$$

$$IOV_{drag} = \frac{|\theta_{drag} - \varphi_{drag}| + |\theta_{drag} - \psi_{drag}| + |\varphi_{drag} - \psi_{drag}|}{3}, \tag{3.1}$$

$$IOV_{xyz} = \frac{|\theta_{xyz} - \varphi_{xyz}| + |\theta_{xyz} - \psi_{xyz}| + |\varphi_{xyz} - \psi_{xyz}|}{3}.$$

Ideally, there should be no variation, so that $IOV = 0$.

Table 1 shows summary statistics, such as the mean and median values, for the IOV for the three methods. The median for the manual IOV ($4.67°$) is marginally better than the expected variation ($5°$) described in Capaldo and Blumenschine (1994) and Draper et al. (2011) but the mean for the manual IOV ($7.08°$) is over $2°$ higher than expected and the standard deviation ($8.48°$) is high. All but seven of the manual IOVs are $< 31°$. The remaining seven are $> 50°$ and could be considered outliers to which the mean and standard deviation are sensitive. However, removing those seven IOVs would not sufficiently reduce these values because the median IOV for the drag method ($2.28°$) and the standard deviation ($3.34°$) are considerably smaller. The xyz method has a consistently smaller IOV compared to the other methods, with a median of $0.001°$, mean $0.006°$, and standard deviation $0.011°$.

Table 1: Summary statistics for angle IOV°

STATISTICS	MANUAL	DRAG	XYZ
N	500	500	500
MIN	0	0.04	0
MEAN	7.08	3.27	0.006
MEDIAN	4.67	2.28	0.001
MAX	73.33	23.15	0.06
SD	8.48	3.34	0.011

Figure 6 shows histograms of the IOV for each method. We see that the IOV for the manual and drag method are rather dispersed with a larger proportion of breaks characterized by larger errors, while the IOV for the xyz method is highly concentrated around very low variabilities. We point out to the reader that the scale of the histogram axes are different in each case. The much smaller range and limited dispersion suggests that the virtual goniometer, regardless of method, outperforms the contact goniometer and the xyz method is far more precise than both the click-and-drag and manual methods.
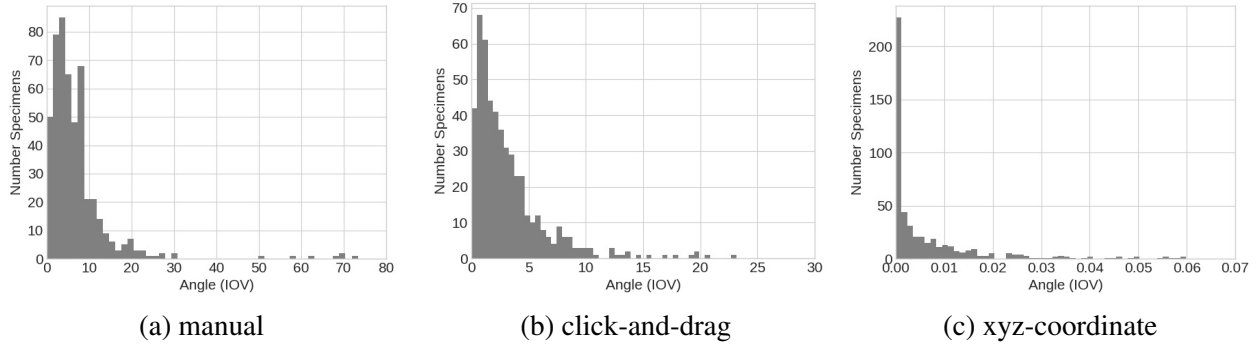
(a) manual   (b) click-and-drag   (c) xyz-coordinate

Figure 6: Histograms of angle IOV$^\circ$ using all three methods (n=500)

3.2 Confidence intervals

We calculated 95% and 99% confidence intervals, the results of which can be found in Table 2 and Figure 7 (Weisberg, 2005). Even with an increase in the confidence level to 99%, the difference is striking — indeed, we have to apply a magnification to observe the xyz-method's confidence interval. The range for the manual method is 2.5 times larger than the click-and-drag method and between 760.5 and 772.9 times larger than the xyz method. The drag method is a little over 300 times larger than the xyz method. The small range IOV of the xyz method indicates the method is exceptionally precise, especially compared to the drag and manual methods.

Table 2: Confidence Intervals

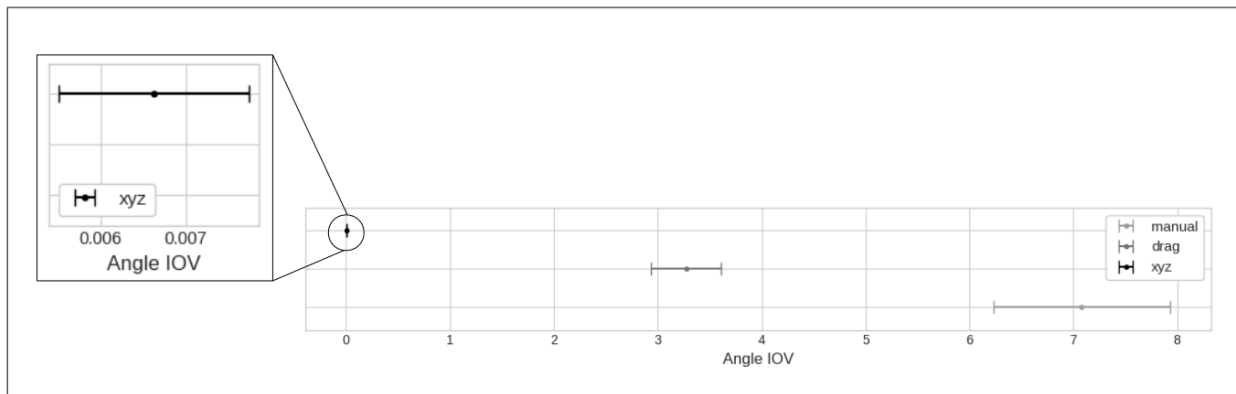| METHOD	MEAN	95% CI			99% CI		
		INTERVAL	RANGE	MEAN ERROR	INTERVAL	RANGE	MEAN ERROR
MANUAL	7.08	(6.2298, 7.9302)	1.7004	±0.74345	(6.0153, 8.1447)	2.1294	±0.62207
DRAG	3.2740	(2.9388, 3.6093)	0.6705	±0.29318	(2.8541, 3.6939)	0.8398	±0.24531
XYZ	0.0066	(0.0055, 0.0077)	0.0022	±0.00097	(0.0052, 0.0080)	0.0028	±0.00082

Figure 7: Plot of the Confidence Intervals for the IOV's (all methods).

We ran Tukey's HSD (honestly significant difference) test with confidence level $alpha = 0.05$ to assess the significance of the difference in means of the IOV scores for each method (Barnette & McLean, 1998).

11

Table 3 shows that the virtual goniometer's drag method is 3.6 degrees ($\pm$ 0.8, 95% C.I.) more consistent than the contact goniometer and the virtual goniometer's xyz method is 7.1 degrees ($\pm$ 0.8, 95% C.I.) more consistent than the contact goniometer (p = 0.001, in both cases). Furthermore, the xyz method is 3.4 degrees ($\pm$ 0.8, 95% C.I.) more consistent than the drag method. The HSD test rejected the null hypothesis that any of the mean IOV values are equal with statistical significance $p < 0.001$. These methods are not equally effective. It is clear that the xyz method is far superior to both the manual and the drag methods.

Table 3: Results of the Tukey HSD Test

GROUP1	GROUP2	MEANDIFF	P-ADJ	LOWER	UPPER	REJECT
DRAG IOV	MAN IOV	3.637	0.001	2.873	4.402	TRUE
XYZ IOV	DRAG IOV	3.436	0.001	2.685	4.187	TRUE
XYZ IOV	MAN IOV	7.073	0.001	6.309	7.838	TRUE

3.3 Summary statistics for the IOV (virtual methods)

Since the center, radius, and mesh point data cannot be collected using the manual method, we compare the drag and xyz methods for the entire sample using calculations of the IOV for each variable. The IOV of the angle measurement for the click-and-drag method features numerous large values and has a standard deviation of 3.7°, varying as much as 28.2° (see Figure 8). The xyz method has a standard deviation of 0.01° with a maximum variation of 0.06°. Most of the xyz IOVs fall below 0.02°. It is clear that the variation in the angle IOV is the result of variation in the location (represented by the center of the patch), the patch's radius, and the number of mesh points in the patch (see Table 4).

Table 4 gives summary statistics for the IOV for the drag and xyz methods based on IOVs for the angle measurement, the number of mesh points in the selected patch, the radius of the patch, and the center location of the patch.

Table 4: Summary statistics for IOV°

| STATISTICS	CLICK-AND-DRAG				XYZ-COORDINATES			
	ANGLE	POINTS	RADIUS	CENTER	ANGLE	POINTS	RADIUS	CENTER
N	537	537	537	537	537	537	537	537
MIN	0.04	2.67	0.01	0	0	0	0	0
MEAN	3.44	270.25	0.52	1.74	0.01	0.39	0.00	0
MEDIAN	2.32	170.67	0.41	1.09	0.00	0.67	0.00	0
MAX	28.20	2250	2.63	30.49	0.06	0.67	0.03	0
SD	3.70	295.06	0.40	2.41	0.01	0.33	0.00	0

3.4 Multiple regression

To better understand how changes in the location, the radius, and the number of points in the patch affect the angle measurement we ran a multiple regression using a log transformation (Weisberg, 2005).

We chose to log-transform the response variable, using the standard natural log, in order to make the data satisfy the assumption of normal error terms (see Figure 9). On the original data the residuals did not
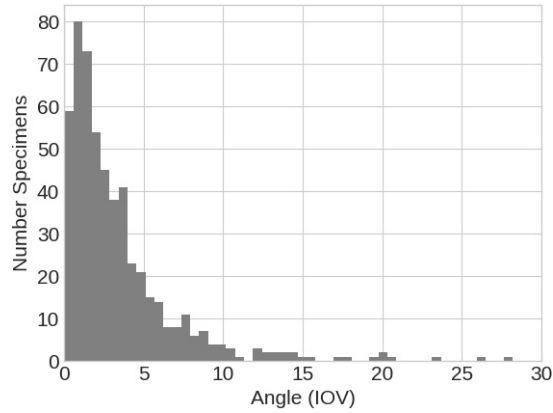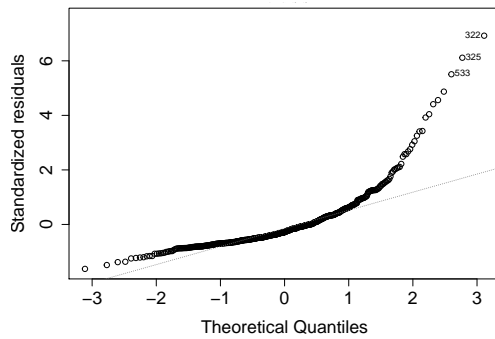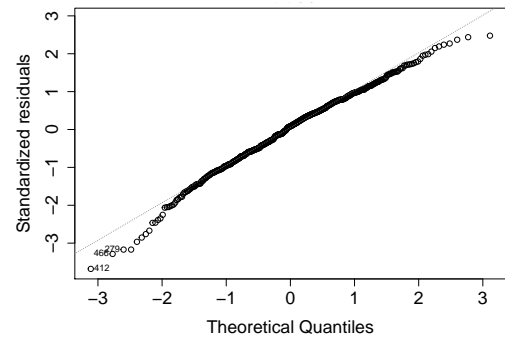
(a) click-and-drag
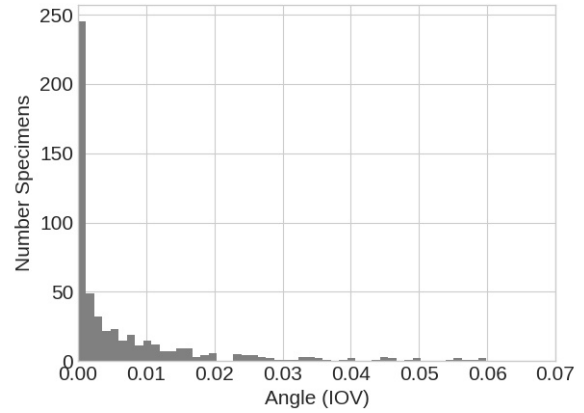
(b) xyz-coordinate

Figure 8: Histograms of angle IOV using all virtual methods (n=537)

(a) Q-Q plot of the original data

(b) Q-Q plot of the log transformed data

Figure 9: Q-Q plots of the residuals(n=537)

follow a normal distribution and were quite skewed (Figure 9a), whereas for the log-transformed data they follow a normal distribution (Figure 9b). Though the residuals do not have an obvious pattern, the Cook's distance plot shows that none of the datapoints are overly influencing the model (Figure 10) (Cook, 1977). None of the observations have high Cook's values which is indicated by the absence of datapoints in the upper or lower right-hand region of the plot.

The results of the multiple regression are presented in Table 5. The p-value of the model's fit is significant ($p = 2.242e - 10$) as are the p-values for each of the explanatory variables indicating that all three variables are influencing the angle IOV. In this model, the radius (estimate = 0.9997) has the biggest impact. The estimate for the mesh points (-0.0009) shows that this has the least impact. The data are highly scattered ($R^2 = 0.0872$) so it would be difficult to predict the angle IOV based on this model. Nonetheless, it is clear that, when replicating measurements, changes in the location, the radius, and the number of mesh points can influence the angle measurement. Having a large number of mesh points within a consistently sized patch with a consistent location is the best method for achieving small IOVs which is why the xyz method is the best option for replication of measurements.
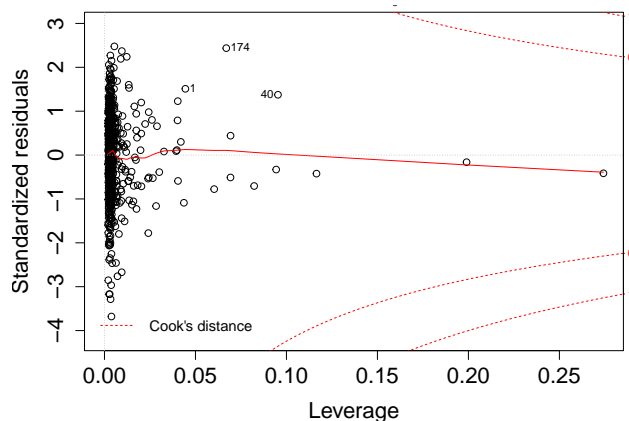
13

Figure 10: Cook's Plot (log-transformed)

Table 5: Multiple Regression Results (n=537)

	Estimate	Std. Error	t value	Pr($>$|t|)
(Intercept)	0.3995	0.0717	5.58	3.94e-08
points IOV	-0.0009	0.0003	-3.6	0.000346
radius IOV	0.9997	0.1993	5.015	7.23e-07
location IOV	0.0573	0.0197	2.909	0.003772

Adjusted R-squared: 0.0872, p-value: 2.242e-10

4 Discussion

Although goniometry plays an increasingly important role in anthropology, as with all measurements made on field data and objects, their overall reliability and hence subsequent inferences, depends on their accuracy, precision, and replicability. And, all of this depends on being able to take the measurement in the first place.

4.1 The physicality of the goniometer

The primary issue with the contact goniometer is the physical interaction between the instrument and its target. The contact goniometer was originally designed to measure the angle between intersecting flat surfaces, such as are found on crystals, and is less well adapted to surfaces with curvature and other features, such as cylindrical long bones, or uneven surfaces that are found on bone and other archaeological materials. The positioning of the goniometer on the object depends on the user, who must ensure that each arm of the instrument lies flat against and perpendicular to the faces being measured. Curvature variations can make this placement challenging.

Because the surfaces are not flat, the angle value depends on the distance from the edge where the measurement is taken on both faces. Stopping 3 mm from the edge of an object or stopping 5 mm from the edge can result in a different angle value. When using the contact goniometer, the distance from the edge generally cannot be controlled. Due to variations in topography among break faces on bone fragments, the arm of the goniometer does not consistently make contact on the surface at the same distance from the edge. When the surface is concave, the goniometer will not rest against the concavity, and will reach across that expanse to the other side.

Conversely, a convex surface prevents the goniometer from reaching the other side of the break. In this

case, measuring an angle that extends across the whole surface is not possible. When the surface is rugose, the goniometer will make contact with the highest point within its reach at that location. Such topographical variations on the surface of the break prevent the goniometer from connecting with most of the surface and it fails to capture a significant amount of intermediate information at a single location.

Furthermore, sometimes the goniometer simply cannot access the necessary location for taking the measurement. It might be blocked by adhering matrix or some other surface feature such as a protuberance on a bone fragment. This can also be a matter of scale in that the goniometer is too large relative to the size and shape of the bone fragment. While a tiny goniometer would, at least in principle, be able to measure smaller or hard-to-reach locations, in practice it would be difficult for the user to handle comfortably and accurately. In some instances, the arms are too long and are blocked by features on the bone fragment or the break on the opposite side. The latter happens when more than half of the circumference of the fragment is present or the angle of interest is acute and descends into the medullary cavity. When measuring notches, Capaldo and Blumenschine (1994) chose to take molds because the goniometer was unable to reach the notch surfaces. However, making molds is often not an option and can even damage the specimen.

One instance in which the virtual goniometer cannot measure a fracture angle is when a sharp natural curve on the bone fragment is close to the fracture edge. This is only an issue when the radius is large. Reducing the radius such that it captures more of the ridge and less of the natural curve rectifies the problem. A smaller radius might be better when measuring angles on bone fragments. Limiting the radius such that it is local to the transition but still sufficiently large to not be hampered by imaging artifacts could provide a more informative measure. As the radius increases in size, it captures changes in the topography. Though the topographical changes on the break surface may also be of interest as it pertains to examining bone fragments, this should remain separate from the angle of transition between faces.

Extremely sharp edges cannot be measured by a contact goniometer (Dibble & Bernard, 1980), because the arms overlap, blocking the location where the specimen is supposed to fit. In some cases, a contact goniometer can only accurately measure angles above $40°$.

Issues arising from the physical interaction of the goniometer and the bone fragment were evident in 20.11% of the breaks in our sample. These breaks were categorized as concave, hinged, surface, edge, or other. If we were physically unable to measure the break, it was categorized as blocked or other (see Table 6). Of the flagged fragments, 108 were measured. Arguably, these measurements do not accurately capture the fracture angle and may not be useful for fracture edge analysis.

As a result of all these physical constraints, measurements taken at an edge angle will be inconsistent (Dibble & Bernard, 1980; Johnson, Dropps, & Yezzi-Woodley, 2019), which calls into question the accuracy of any comparative studies. Another disadvantage of the contact goniometer is the amount of time required to take each manual angle measurement. As a result, its application to assemblages that number in the thousands can become impractical if not impossible to complete. Thus time and physical constraints may make it impossible to capture sufficiently representative data for detailed and rigorous studies.

4.2   Precision

The ability to take a measurement does not ensure its appropriateness. Assuming the angle can be taken and is appropriate to take, it also needs to be precise enough that it is useful. When measuring platform angles of notch molds using the contact goniometer, Capaldo and Blumenschine (1994) expected measurements to vary up to $5°$, which may well not be precise enough. For example, Alcántara-García et al. (2006) state that, in general, carnivores produce fracture angles between $80°$ and $110°$, whereas fracture angles on bones broken by percussion will be $< 80°$ and $> 110°$, offering a $30°$ range for assigning carnivores as the actors responsible for breakage. If one chooses a more stringent approach where fracture angles between $85°$ and

Table 6: Categorized and uncategorized breaks

Category	Count	Percent
ALL BREAKS	537	100%
NOT CATEGORIZED	392	73%
CATEGORIZED	145	27%
BLOCKED[1]	34	6.33%
CONCAVE	70	13.04%
HINGE	21	3.91%
SURFACE[2]	4	0.74%
EDGE[3]	11	2.05%
OTHER	5	0.93%
CATEGORIZED (MEASURED)	108	20.11%
CATEGORIZED (NOT MEASURED)	37	6.89%

[1] THE GONIOMETER ARM CANNOT MAKE CONTACT WITH THE DESIRED SURFACES
[2] THE SURFACE IS RUGOSE AND HAS A LOT TOPOGRAPHICAL RELIEF
[3] THE EDGE IS ROUNDED OR HINDERS EFFORTS TO MEASURE

95° (Pickering et al., 2005), generally labeled as dry breaks, are excluded, this diminishes the ranges for carnivores to 5° and 15°. These are small windows for such a large error range. In fact, one of those ranges is equal to the expected error range. In a test that looked at the reliability of the contact goniometer when measuring knee angle in a medical context, Draper et al. (2011) noted that in order to remain within an error range of 5°, the location of the measurement had to be within 2mm of the actual center of the patella. Not only is this equal to the error range presented by Capaldo and Blumenschine, these results also highlight the importance of the location where the measurement is taken since, in the case of bone fragments, the angle values can vary across the width of the break as well as along the length of the break edge.

Similarly, problems with precision in angle measurements are also recognized in the analysis of other archaeological materials, such as lithic artifacts. A number of studies have been published criticizing the contact goniometer's application in this field, beginning with Dibble and Bernard (1980)'s demonstration of the large inter-observer variation of 16.6° for angles between 25° and 75°. Others, such as Andrefsky Jr (1998, 89-92), Cochrane (2003), Dibble and Whittaker (1981), Gnaden and Holdaway (2000), and Odell (2012) have focused on the difficulties of standardizing the physical placement of the contact goniometer along different stone edges and surfaces, akin to our observations above. Regardless, however, because angles are known to be behaviorally significant for flake production (platform angles: Dibble, 1997; Magnani, Rezek, Lin, Chan, & Dibble, 2014; Režek et al., 2018; Scerri et al., 2016; Tostevin, 2003) and lithic tool function (cutting edge angles: Key & Lycett, 2015; Wilkins et al., 2017), lithic analysts have endeavored to surmount the precision problems in two ways. First, they typically apply different measurement tools to different angles. Specifically, the caliper-goniometer of Dibble and Bernard (1980) is applied to acute angles < 40° (suitable for cutting edge and retouch edge angles) while the contact goniometer is applied to platform angles which are harder to measure with the caliper-goniometer and which are consistently over the 40° threshold for where Dibble & Bernard's data show the two devices to be equally reliable. Other analysts, however, do not specify this shift in tools between different angle ranges, reporting only the use of the contact goniometer for both types of edges (e.g. Pargeter & de la Peña, 2017; Wilkins et al., 2017). Others focus on intra-observer error by choosing to measure lithic angles three times and to use the mean of the observations if the latter are roughly similar to each other (e.g. Hovers, 2009, for interior platform angles).

The second way in which lithic analysts endeavor to control for inter-observer variation is to devote abundant time and effort in training analysts as part of a team, such that all members of the team have spent

many hours in one-on-one instruction on how to use the contact goniometer on over 200 lithic artifacts, representing the range of possible problematic cases for the device. The best example of this approach is the team created by the late Harold Dibble, which resulted recently in the large, multi-authored study of over 18,000 artifacts from 81 assemblages spanning about 2 million years, measured by over a dozen analysts over many decades (Režek et al., 2018). Because all of the authors were trained by the same individual(s), one can have confidence that measurements were taken similarly and problematic pieces were not included in their angle data. Yet this approach has a significant limitation for a wider scientific application: non-team researchers cannot add to their data, even though it is freely available, without risking an unknown increase in inter-observer error. That being said, it is unknown what the actual inter-observer variation is in the Režek et al. (2018) teams' data, although it is certainly better than the 16.6° reported by Dibble and Bernard (1980). Re-imagining their results, however, as the meta-analysis of data from 81 independent researchers working on one assemblage each (rather than a unified study), it would not be inappropriate to worry about the possible combination of the known 16.6° inter-observer variation from Dibble and Bernard (1980) and the intra-observer variation of 5° in edge angles demonstrated by the present study. When looking at their EPA median data (Režek et al., 2018, Figure 1, p. 629), most of which fall within a range of about 65° to 95°, a possible error range of 16° would have a substantial impact on data distribution and subsequent interpretations. While this meta-analysis scenario is hypothetical, such studies are in fact the goal of open access data science and we raise the issue of precision to encourage researchers to be cognizant of how certain variables are measured and any subsequent comparability issues resulting from those methods. All of these examples, from faunal as well as lithic analysis, suggest that a simpler and demonstrably more replicable tool such as the virtual goniometer is preferable for future research across multiple artifact types.

4.3   Experimental replication

Determining an exact location for a measurement makes it easier to replicate. When measuring fracture angles on bone fragments, some researchers measure at the midpoint along the ridge (Coil et al., 2017; Dibble & Bernard, 1980; Pickering et al., 2005) whereas others use the most extreme angle (Moclán et al., 2019). Capaldo and Blumenschine (1994) define a midpoint, but, it only applies to notches and cannot be extrapolated to all fracture edges because it depends on features that are specific to notches, specifically inflection points.

Finding the midpoint on a break edge is less clear. The midpoint on a bone fracture edge as described by Pickering et al. (2005) and Coil et al. (2017) is likely the approximate midpoint as opposed to an exact midpoint. The break edge on each break face on a bone fragment can be viewed as a contour. Establishing the endpoints for the contour is the first challenge. Typically, the full length of the break face does not terminate at the same location that the ridge between the periosteal surface and break surface terminates. Once a decision is made as to where the endpoints of the contour will be located, then one can decide where to take the angle measurement. In regard to the midpoint, one could choose the midpoint of the Euclidean distance between the two end points of that contour or one could choose the midpoint of the full contour length. In either case, finding the exact midpoint on the physical object is challenging, if not impossible, and time consuming. For example, calipers could be used in the case of the Euclidean midpoint. However, this would require a consistent orientation of the specimen in relationship to the calipers. Finding the most extreme angle would require approximation as well unless one measures many angles along the edge in order to identify which one is the most extreme. This requires that the angle can be measured at all locations and the goniometer is not blocked by features on the specimen at any location. When using 3D models, specific points can be chosen as endpoints for the contour. The Euclidean distance or the contour length can be calculated and a midpoint can be consistently defined and extracted.

Whether choosing the most extreme measure or the midpoint, taking a measurement at a single location

is anthropologically arbitrary. Bones are not limited to one instance of fragmentation. If the objective is to identify the first actor of breakage, then the extrema or midpoints will not be comparable among specimens if additional fragmentation takes place. An alternative approach, that could provide a richer dataset, would be to take multiple measurements along the edge of the break which can be done quickly using the virtual goniometer. In any case, experimental replication is efficient and simple using the virtual goniometer because it outputs the data required to precisely replicate each measurement.

4.4 Accuracy

The accuracy of the virtual goniometer depends on many non-algorithmic factors, including the resolution of the 3D model, the size of the object being analyzed, the area of interest on the object, the degree to which there are topographical changes on the surface, and the scale at which the object is to be analyzed. Thus, it is not possible or appropriate to recommend an optimal scan resolution that will work for all anthropological studies.

In order to explore how the resolution of the scan affects the accuracy of the angle measurement, we tested the virtual goniometer on synthetic meshes consisting of two planes intersecting at a known angle. We varied the mesh resolution and the angle, and subjected the virtual goniometer to extreme cases that are generally beyond the range one would use in practice (i.e., very low resolutions).

The vertices of the synthetic meshes were chosen as independent and uniformly distributed random variables, and the mesh was generated with a Delaunay triangulation (Cheng, Dey, & Shewchuk, 2012). The two intersecting planes are identical, with side length of 1 along the break edge and 0.5 perpendicular to the break edge. We generated synthetic meshes with angles ranging from $1°$ to $179°$ in increments of $10°$ with additional meshes at $1°$, $5°$, $175°$, and $179°$ angles. For each angle we varied the number of points (i.e., the resolution) in the mesh from 50 up to $10,000$ points, and we generated 30 synthetic meshes for each combination of angle and resolution. For each mesh, we used the virtual goniometer to measure the angle at the center of the break with a radius of $r = 0.4$, and segmentation tuning parameter $\lambda = 0$ (since the meshes do not have noise). Because this was an automated process, there is no graphical user interface and the tuning parameter cannot be adjusted for individual measurements to improve the segmentation between planes. Therefore, the segmentation is completely reliant on resolution and some level of variation is expected among measurements.

The number of points in the $r = 0.4$ radius patch ranged from 20 to $5,186$ in the experiment, resulting in a sample of $15,702$ synthetic meshes. We removed all observations where the number of points in the $r = 0.4$ radius patch was < 20. The virtual goniometer requires a sufficient number of points in order to produce meaningful results. Nevertheless, we find that the virtual goniometer performs well even at quite low resolutions; the results of our experiment are displayed in Figure 11. The acceptable threshold for variation in the measurement will depend on the context of the research.

4.5 Future research

The next logical step is to start using the tool to address specific anthropological questions. Though the virtual goniometer is a broadly applicable tool that can accurately and consistently measure angles, its usefulness within specific artifact classes – including but not limited to pottery, stone, metal, and glass – needs to be tested and offers areas for future research. Questions to explore include where and how to take measurements and appropriate scan resolutions for specific artifact types and research questions. To that end, we are currently exploring application of the virtual goniometer on lithic artifacts and continuing to test its utility for fracture angle analysis on bone fragments.
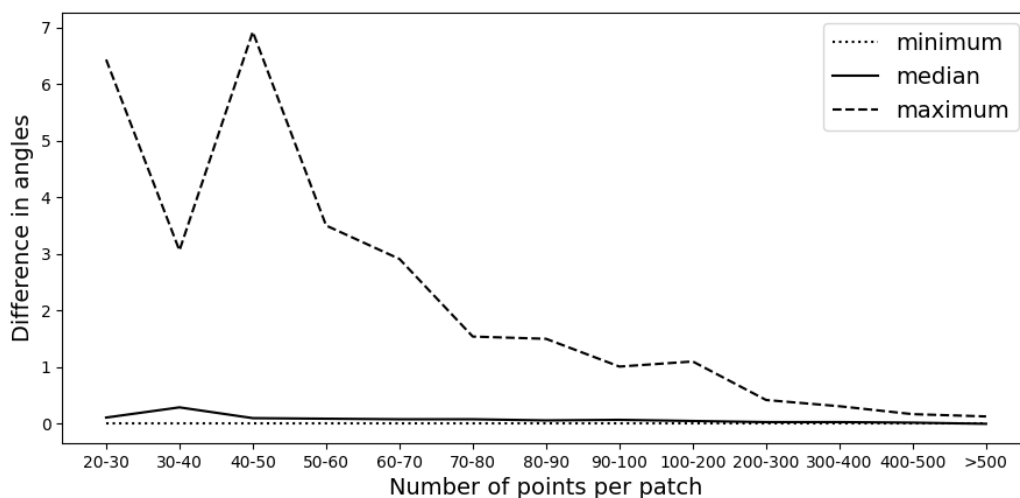
Figure 11: Plot of the difference between the true and measured angles at different resolutions.

5 Conclusion

Scanning objects is becoming mainstream in anthropology and the virtual goniometer is a logical next step that is easy to integrate into research that uses 3D models.

The purpose of this project was to introduce the virtual goniometer and demonstrate its capabilities. We have demonstrated that the precision and accuracy of the virtual goniometer far surpasses the capabilities of the contact goniometer. Additionally, the virtual goniometer automatically extracts the segmentation parameter, a measure of the goodness of fit, the measurement location, the radius of the selected area, and the number of mesh points used in the calculation, none of which can be garnered by using the contact goniometer. By providing numerical metadata for the measurement and 3D visualizations, inter- and intra-observer discrepancies can be easily identified and measurements can be replicated with precision. The virtual goniometer provides flexibility that allows the user to adjust parameters and choose an approach that is dependable and useful. The virtual goniometer resolves the inherent limitations of the contact goniometer without limiting its application to specific edge morphologies or artifact classes, as with extant 3D angle measurement tools. Further, it gives researchers a tangible and consistent way to discuss how best to employ goniometry to address questions in anthropology.

improve this paper.

References

Adamopoulos, E., Rinaudo, F., & Ardissono, L. (2021). A critical comparison of 3d digitization techniques for heritage objects. *ISPRS International Journal of Geo-Information*, *10*(1), 10.

Alcántara-García, V., Egido, R. B., del Pino, J. M. B., Ruiz, A. B. C., Vidal, A. I. E., Aparicio, Á. F., . . . others (2006). Determinación de procesos de fractura sobre huesos frescos: un sistema de análisis de los ángulos de los planos de fracturación como discriminador de agentes bióticos. *Trabajos de prehistoria*, *63*(1), 37–45.

Andrefsky Jr, W. (1998). *Lithics*. Cambridge University Press.

Archer, W., Gunz, P., van Niekerk, K. L., Henshilwood, C. S., & McPherron, S. P. (2015). Diachronic change within the still bay at blombos cave, south africa. *PLoS One*, *10*(7), e0132428.

Archer, W., Pop, C. M., Gunz, P., & McPherron, S. P. (2016). What is still bay? human biogeography and bifacial point variability. *Journal of Human Evolution*, *97*, 58–72.

Barnes, A. S. (1939). The differences between natural and human flaking on prehistoric flint implements. *American Anthropologist*, *41*(1), 99–112.

Barnette, J. J., & McLean, J. E. (1998). The tukey honestly significant difference procedure and its control of the type i error-rate.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

Burchard, U. (1998). History of the development of the crystallographic goniometer. *Mineralogical Record*, *29*(6), 517.

Capaldo, S. D., & Blumenschine, R. J. (1994). A quantitative diagnosis of notches made by hammerstone percussion and carnivore gnawing on bovid long bones. *American Antiquity*, *59*(4), 724–748.

Cheng, S.-W., Dey, T. K., & Shewchuk, J. (2012). *Delaunay mesh generation*. CRC Press.

Cochrane, G. W. (2003). On the measurement and analysis of platform angles. *Lithic Technology*, *28*(1), 13–25.

Coil, R., Tappen, M., & Yezzi-Woodley, K. (2017). New analytical methods for comparing bone fracture angles: A controlled study of hammerstone and hyena (*Crocuta crocuta*) long bone breakage. *Archaeometry*, *59*(5), 900–917.

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, *19*(1), 15–18.

Das, A. J., Murmann, D. C., Cohrn, K., & Raskar, R. (2017). A method for rapid 3d scanning and replication of large paleontological specimens. *PloS one*, *12*(7), e0179264.

De Juana, S., & Dominguez-Rodrigo, M. (2011). Testing analogical taphonomic signatures in bone breaking: A comparison between hammerstone-broken equid and bovid bones. *Archaeometry*, *53*(5), 996–1011.

Dibble, H. L. (1997). Platform variability and flake morphology: a comparison of experimental and archaeological data and implications for interpreting prehistoric lithic technological strategies. *Lithic technology*, *22*(2), 150–170.

Dibble, H. L., & Bernard, M. C. (1980). A comparative study of basic edge angle measurement techniques. *American Antiquity*, *45*(4), 857–865.

Dibble, H. L., & Whittaker, J. C. (1981). New experimental evidence on the relation between percussion flaking and flake variation. *Journal of Archaeological Science*, *8*(3), 283–296.

Domínguez-Rodrigo, M., & Barba, R. (2006). New estimates of tooth mark and percussion mark frequencies at the FLK Zinj site: the carnivore-hominid-carnivore hypothesis falsified. *Journal of Human Evolution*, *50*(2), 170–194.

Draper, C. E., Chew, K. T., Wang, R., Jennings, F., Gold, G. E., & Fredericson, M. (2011). Comparison of quadriceps angle measurements using short-arm and long-arm goniometers: correlation with mri. *Pm&r*, *3*(2), 111–116.

Gifford-Gonzalez, D. (1989). Ethnographic analogues for interpreting modified bones: some cases from east africa. *Bone modification*, 179–246.

Gnaden, D., & Holdaway, S. (2000). Understanding observer variation when recording stone artifacts. *American Antiquity*, 739–747.

Gould, R. A., Koster, D. A., & Sontz, A. H. (1971). The lithic assemblage of the western desert aborigines of australia. *American Antiquity*, *36*(2), 149–169.

Grosman, L., Goldsmith, Y., & Smilansky, U. (2011). Morphological analysis of nahal zihor handaxes: a chronological perspective. *PaleoAnthropology*, *2011*, 203–215.

Grosman, L., Karasik, A., Harush, O., & Smilansky, U. (2014). Archaeology in three dimensions: Computer-based methods in archaeological research. *Journal of Eastern Mediterranean Archaeology and Heritage Studies*, *2*(1), 48–64.

Grosman, L., Smikt, O., & Smilansky, U. (2008). On the application of 3-d scanning technology for the documentation and typology of lithic artifacts. *Journal of Archaeological Science*, *35*(12), 3101–3110.

Han, S., & Boutin, M. (2014, Nov). *RP1D (Clustering Method Using 1D Random Projections)*. (Code: https://purr.purdue.edu/publications/1729/1)

Haynes, G. (1983). Frequencies of spiral and green-bone fractures on ungulate limb bones in modern surface assemblages. *American antiquity*, 102–114.

Hovers, E. (2009). *The lithic assemblages of qafzeh cave*. Oxford University Press.

Johnson, K., Dropps, K., & Yezzi-Woodley, K. (2019). Testing for inter- and intra-observer error in bone fracture angle measurement data. In *Poster presented at annual meeting of the paleoanthropology society. albuquerque, new mexico.*

Key, A. J., & Lycett, S. J. (2015). Edge angle as a variably influential factor in flake cutting efficiency: an experimental investigation of its relationship with tool size and loading. *Archaeometry*, *57*(5), 911–927.

Kuhn, S. L. (1990). A geometric index of reduction for unifacial stone tools. *Journal of Archaeological Science*, *17*(5), 583–593.

Lerman, G., & Maunu, T. (2018). An overview of robust subspace recovery. *Proceedings of the IEEE*, *106*(8), 1380–1410.

Magnani, M., Rezek, Z., Lin, S. C., Chan, A., & Dibble, H. L. (2014). Flake variation in relation to the application of force. *Journal of Archaeological Science*, *46*, 37–49.

Maté-González, M. Á., Aramendi, J., González-Aguilera, D., & Yravedra, J. (2017). Statisti-

cal comparison between low-cost methods for 3d characterization of cut-marks on bones. *Remote Sensing*, *9*(9), 873.

Moclán, A., Domínguez-Rodrigo, M., & Yravedra, J. (2019). Classifying agency in bone breakage: an experimental analysis of fracture planes to differentiate between hominin and carnivore dynamic and static loading using machine learning (ml) algorithms. *Archaeological and Anthropological Sciences*, *11*(9), 4663–4680.

Moclán, A., Huguet, R., Márquez, B., Laplana, C., Arsuaga, J. L., Pérez-González, A., & Baquedano, E. (2020). Identifying the bone-breaker at the navalmaíllo rock shelter (pinilla del valle, madrid) using machine learning algorithms. *Archaeological and Anthropological Sciences*, *12*(2), 46.

Morales, J. I., Lorenzo, C., & Vergès, J. M. (2015). Measuring retouch intensity in lithic tools: a new proposal using 3d scan data. *Journal of Archaeological Method and Theory*, *22*(2), 543–558.

Nigst, P. R. (2012). *The early upper palaeolithic of the middle danube region*. Leiden University Press.

Odell, G. H. (2012). *Lithic analysis*. Springer Science & Business Media.

Pargeter, J., & de la Peña, P. (2017). Milky quartz bipolar reduction and lithic miniaturization: Experimental results and archaeological implications. *Journal of Field Archaeology*, *42*(6), 551–565.

Pickering, T. R., Domínguez-Rodrigo, M., Egeland, C. P., & Brain, C. (2005). The contribution of limb bone fracture patterns to reconstructing early hominid behaviour at swartkrans cave (south africa): archaeological application of a new analytical method. *International Journal of Osteoarchaeology*, *15*(4), 247–260.

Pickering, T. R., & Egeland, C. P. (2006). Experimental patterns of hammerstone percussion damage on bones: implications for inferences of carcass processing by humans. *Journal of archaeological Science*, *33*(4), 459–469.

Pop, C. (2019). *Lithics3d*. Retrieved from https://github.com/cornelmpop/Lithics3D (v0.4.2, April 21 2019)

Porter, S. T., Huber, N., Hoyer, C., & Floss, H. (2016). Portable and low-cost solutions to the imaging of paleolithic art objects: A comparison of photogrammetry and reflectance transformation imaging. *Journal of Archaeological Science: Reports*, *10*, 859–863.

Porter, S. T., Roussel, M., & Soressi, M. (2016). A simple photogrammetry rig for the reliable creation of 3d artifact models in the field: lithic examples from the early upper paleolithic sequence of les cottés (france). *Advances in Archaeological Practice*, *4*(1), 71–86.

Režek, Ž., Dibble, H. L., McPherron, S. P., Braun, D. R., & Lin, S. C. (2018). Two million years of flaking stone and the evolutionary efficiency of stone tool technology. *Nature ecology & evolution*, *2*(4), 628–633.

Sapirstein, P., & Murray, S. (2017). Establishing best practices for photogrammetric recording during archaeological fieldwork. *Journal of Field Archaeology*, *42*(4), 337–350.

Scerri, E. M., Gravina, B., Blinkhorn, J., & Delagnes, A. (2016). Can lithic attribute analyses identify discrete reduction trajectories? A quantitative study using refitted lithic sets. *Journal of Archaeological Method and Theory*, *23*(2), 669–691.

Tostevin, G. B. (2003). Attribute analysis of the lithic technologies of stránská skála ii-iii in their regional and inter-regional context: Origins of the upper paleolithic in the brno basin. In *Stránská skála: Origins of the upper paleolithic in the brno basin* (pp. 77–118). Peabody

Museum Press, Harvard University.

Valletta, F., Smilansky, U., Goring-Morris, A. N., & Grosman, L. (2020). On measuring the mean edge angle of lithic tools based on 3-d models–a case study from the southern Levantine Epipalaeolithic. *Archaeological and Anthropological Sciences*, *12*(2), 49.

Villa, P., & Mahieu, E. (1991). Breakage patterns of human long bones. *Journal of human evolution*, *21*(1), 27–48.

Weisberg, S. (2005). *Applied linear regression* (Vol. 528). John Wiley & Sons.

Weiss, M. (2020). The lichtenberg keilmesser-it's all about the angle. *Plos one*, *15*(10), e0239718.

Weiss, M., Lauer, T., Wimmer, R., & Pop, C. M. (2018). The variability of the keilmesser-concept: a case study from central germany. *Journal of Paleolithic Archaeology*, *1*(3), 202–246.

Wilkins, J., Brown, K. S., Oestmo, S., Pereira, T., Ranhorn, K. L., Schoville, B. J., & Marean, C. W. (2017). Lithic technological responses to late pleistocene glacial cycling at pinnacle point site 5-6, south africa. *PloS one*, *12*(3), e0174051.

Wilmsen, E. N. (1968). Functional analysis of flaked stone artifacts. *American Antiquity*, *33*(2), 156–161.

Yellamraju, T., & Boutin, M. (2018). Clusterability and clustering of images and other "real" high-dimensional data. *IEEE Transactions on Image Processing*, *27*(4), 1927–1938.

Zhang, T., Szlam, A., Wang, Y., & Lerman, G. (2010). Randomized hybrid linear modeling by local best-fit flats. In *2010 ieee computer society conference on computer vision and pattern recognition* (pp. 1927–1934).