Journal of Archaeological Science: Reports 17 (2018) 597–604

Contents lists available at ScienceDirect

Journal of Archaeological Science: Reports

journal homepage: www.elsevier.com/locate/jasrep

# An experimental investigation of the effects of post-depositional damage on current quantitative use-wear methods

J. Jeffrey Werner*

*Department of Anthropology, University of Alberta, Edmonton, Canada*

ARTICLE INFO

*Keywords:*
Use-wear
Quantitative methods
Microscopy
Post-depositional damage
Experimental archaeology

ABSTRACT

This paper describes a new method of estimating stone tool function using a laser-scanning confocal microscope. The primary benefit of the quantitative approach used here, and others like it, is that it increases reproducibility while reducing subjectivity and inter/intra-observer error. Although the use of laser-scanning confocal microscopy in functional analyses is not novel, the method presented here was "stress tested" using an experimental assemblage to explore the impact of post-depositional damage on its results. The findings reveal that this method is still viable with lightly and moderately damaged specimens. It is therefore likely that methods like this one are no more vulnerable to post-depositional damage than conventional use-wear methods and can confidently be applied to archaeological specimens.

## 1. Introduction

Use-wear analysis presents archaeologists with a powerful tool with which to directly assess the function of artifacts without relying on morpho-functional arguments or by drawing analogies to modern tools. For the most part, these methods rely on a visual comparison of the damage formed by use on experimentally replicated tools and archaeological specimens. The origins of the field dates back to the 1960s, with interest among western archaeologists first peaking around 1980 (Keeley, 1980; Semenov, 1964; Stemp et al., 2016). Early on, however, researchers began to voice concerns. Critics claimed that findings were difficult to interpret, and that there was no standardized language with which to describe wear traces (Grace, 1996). Blind-testing also demonstrated that the success of conventional methods are heavily dependent on the training and experience of individual analysts, causing overall accuracy to vary widely across studies (Evans, 2014).

Although there has been significant progress since the 1960s, the field continues to struggle with four key issues (Evans and Donahue, 2008). First, the formation of use-wear is not yet completely understood, though detailed research by authors such as Stemp et al. (2015) and Ollé and Vergès (2014), among others, has been extremely helpful in this regard. Second, there has arguably been an overemphasis on the study of flint/chert assemblages leaving use-wear on non-flint toolstones under-examined. This problem is also being approached by more comprehensive study of non-flint artifacts (Conte et al., 2015; Fernández-Marchena and Ollé, 2016). Third, many aspects of the burial environment are known to be capable of interfering with the interpretability of use-traces (Shea and Klenck, 1993; Venditti et al., 2016). Research suggests that natural transformations of stone artifacts may be mistaken for use-wear and can severely limit the information available from heavily altered pieces. The most common procedure for addressing the issue of post-depositional wear is to exclude artifacts or assemblages from analysis that are suspected to have been naturally damaged (Burroni et al., 2002). This practice remains an imperfect but potentially unavoidable solution. Lastly, and perhaps most challenging, the majority of use-wear analyses remain inherently subjective, difficult to reproduce and independently verify. Ideas vary as to how best to address this final challenge. Some researchers have proposed that analysts receive more intensive training within the framework of existing methods and argue for the creation of larger and more complete reference collections (Rots and Plisson, 2014). Alternately, others have looked to quantitative methods to improve accuracy and objectivity (Dumont, 1982; Evans and Donahue, 2008; González-Urquijo and Ibáñez-Estévez, 2003; Ibáñez et al., 2014; Macdonald, 2014; Macdonald and Evans, 2014; Stemp, 2014; Stemp et al., 2013, 2015, 2016).

Use-wear quantification refers to a field of related techniques that seek to mathematically characterize and differentiate damage on stone tools. These approaches allow data to be statistically, rather than visually, compared to an experimental reference to determine the likely source, or sources of damage. While some scholars have focused on image analysis and damage distribution patterning (Bird et al., 2007; González-Urquijo and Ibáñez-Estévez, 2003; Schoville, 2010; Wilkins and Schoville, 2016), most recent emphasis has been on measuring the profile or areal roughness of stone tools using high-resolution scanning

---

* 116 St and 85 Ave, T6G 2R3, Canada.
*E-mail address:* jwerner@ualberta.ca.

https://doi.org/10.1016/j.jasrep.2017.12.008
Received 6 September 2017; Received in revised form 22 November 2017; Accepted 8 December 2017

equipment borrowed from fields such as engineering and materials science (Evans and Donahue, 2008; Macdonald, 2014; Stemp, 2014; Stevens et al., 2010). One of the primary benefits of these approaches is that the results can be expressed probabilistically, allowing for better transparency and confidence in the reporting of findings.

Although use-wear quantification offers compelling advantages over conventional techniques, the field is still developing, and some of the problems inherent to conventional use-wear analysis are shared by both approaches. The purpose of this article is not to fully solve for all of these problems, but to nudge quantitative methods closer to widespread adoption by studying the impacts of some of these issues. In particular, this experiment tests the sensitivity of measurements obtained with 3D scanning equipment to the distorting effects of post-depositional damage. Although the impact of the burial environment on artifacts of all kinds has been well studied (Lyman, 1994; Marreiros et al., 2015; McBride and Mercer, 2012; Shea and Klenck, 1993), the impact of trampling, patination, movement through the sediment, etc. on 3D scanning data has yet to be established through experimentation. Because of the precise nature of these measurements, it is conceivable that even minor amounts of post-depositional alteration could render these methods invalid. As a benchmark, Shea and Klenck (1993) observed that even short intervals of trampling (15 min) can significantly reduce the accuracy of conventional use-wear results. To investigate the effects of these processes on quantitative data the method detailed in this article was "stress-tested" by exposing the samples to increasing degrees of simulated post-depositional damage. The primary questions of interest were:

1. At what degree does post-depositional damage have a significant effect on the accuracy of scanning metrics?
2. Does additional damage correlate linearly with diminished quantitative use-wear accuracy?
3. Which types of use-wear are most likely to be obscured by damage introduced by post-depositional processes?
4. Is damage emulative of any particular use-traces?

## 2. Method

The methods used were loosely based on Shea and Klenck (1993), who investigated the effects of post-depositional damage on conventional use-wear interpretability. To test these effects, they assessed the ability of a skilled analyst (Shea) to provide accurate functional estimates for specimens that had been damaged by cumulative 15 min increments of trampling.

For the purpose of this experiment, an experimental assemblage of ten unretouched dacite flakes was manufactured. Dacite is a glassy, usually grey, volcanic stone that has similar flaking properties to obsidian (Fig. 1). All flakes in this study were produced using a quartzite hammer-stone by the author. With the exception of two unused control flakes, each tool was used in a sawing motion to process one of four materials for 40 min (two flakes per material type). The contact materials included in this experiment were wood covered in bark, dry antler, dry hide and dry grass stems and leaves. Once the flakes were used, they were carefully cleaned to prevent the introduction of spurious use-wear traces. The experimental assemblages was washed by hand with a grit-free detergent, avoiding brushes or other abrasive materials (Stemp, 2014). They were then soaked in a dilute HCL solution and NaOH respectively for 10 min each to remove any lingering residues or particles (Evans and Macdonald, 2011).

Because of its long working distance, worn sections of each piece were initially identified with a Nikon Eclipse LV150 optical microscope using the 20 × and 50 × objective lenses. These sections were marked with ink so they could be identified and re-scanned later. This preparatory step made it possible to directly trace the changes in surface roughness of a single segment of the tool. Three such areas were located on the ventral side of each piece, generally within 50–100 μm of the working margin. Polished areas were identified and measured although chipping, rounding, as well as series of striations sub-parallel to one another and to the direction of tool motion were also observed (Fig. 2). These wear signatures occurred most commonly on raised sections of the microtopography and are consistent with the sawing motion used by the experimenter. These locations were chosen randomly in the case of unused specimens.

After initial inspection, the specimens were scanned in the three places marked earlier using an Olympus laser-scanning confocal microscope (LSCM) (LEXT OLS3000) located in the materials sciences department at the University of Alberta. The device reconstructs the surface of the tool using light reflected back from the specimen, collected at varying vertical increments. Its operation and anatomy are well described in Evans and Donahue (2008). A LSCM was chosen for this experiment for several reasons. Most importantly, its ability to differentiate contact materials has already been established by prior research (Evans and Donahue, 2008; Ibáñez et al., 2014; Macdonald and Evans, 2014). Furthermore, the instrument is capable of capturing areal as well as profile measurements and each scan can be acquired within a minute or two. The wavelength of the laser light used by this particular system is 408 nm, the horizontal resolution is 0.12 μm and the vertical resolution is 0.10 μm. Scans were taken using the 50 × objective lens (NA = 0.95, WD = 0.3 mm), as recommended by the operator's manual, and each scan took under 2 min to complete. The "fine" setting was selected to enable the capture of detailed measurements. This setting priorities high fidelity data capture as opposed to speed.

LEXT OLS4100 software was used to process the 3D renderings and to take measurements of their surfaces. The "surface correction" tool was used to remove differences in the inclination of the samples. A filter was also used to separate the surface texture of the pieces into roughness and waviness data sets. Waviness refers to longer wave-length surface irregularities, while roughness refers to irregularities with a shorter wave-length and a greater frequency. The distinction between these data sets is based on several agreed upon cut-off points, including 8 μm, 12 μm, and 25 μm. In this case, a cut-off of 8 μm was chosen to isolate and quantify the roughness profiles of the flakes as per previous experimentation (Ibáñez et al., 2014). Measurements were taken at two different areas sizes as per Evans and Donahue (2008), specifically from 10 μm and 100 μm squares (Fig. 3). By varying the capture size of the measurements it is hypothetically possible to account for features of different sizes. Fifty measurements (from here on referred to as cases) were captured for each zone, totaling 150 cases for each specimen and 1500 cases overall. These capture areas were selected so that only polished areas were contained within them, excluding non-worn surfaces.

Five different roughness parameters were captured for each case at two different scales. A roughness parameter is a mathematical formula that describes an aspect of the surface. Because of the complexity of lithic microtopography, it is unlikely that a single parameter is capable of fully describing a surface and it is currently not clear which parameter, or combination of parameters, are best suited to the task of discriminating between worn surfaces (but see Watson and Gleason (2016) for bone tools). Parameters that have proven effective, either individually or in combination, include the root square mean height of the surface (Sq), the arithmetic mean height of the surface (Sa), the maximum peak height (Sp), the maximum valley depth (Sv), the summed maximum peak height and valley depth (Sz), relative area: the ratio between the height of a material at a given threshold and the evaluation area (Sal) and the extreme peak height (Sdc). In this study, the parameters Sq, Sp, Sv, Sz and Sa (all the parameters available for measurement in the program) were used to create a multinomial logistic regression model in IBM SPSS v24. Multinomial logistic regression is a statistical technique used to predict group membership from a set of continuous and/or categorical variables. In this way, it is similar to discriminant function analysis, but has fewer assumptions and can accommodate a greater variety of data types. All variables were log