DETERMINING HOMINID HANDEDNESS IN LITHIC DEBITAGE: A REVIEW OF CURRENT METHODOLOGIES

LANA RUCK[1], DOUGLAS C. BROADFIELD[2], AND CLIFFORD T. BROWN[1]

[1] *Department of Anthropology, Florida Atlantic University, 777 Glades Rd., Boca Raton, FL, USA*
[2] *Department of Cell Biology, Miller School of Medicine, University of Miami, 1600 Northwest 10th Avenue #1140, Miami, FL, USA*

*Handedness is inextricably linked to brain lateralization and language in humans, and identifying handedness in the paleo-archaeological record is important for understanding hominid cognitive evolution. This study reports on experiments for identifying knapper handedness in lithic debitage using three previously established methods: Toth (1985), Rugg and Mullane (2001), and Bargalló and Mosquera (2013). A blind study was conducted on lithic debitage (*n = 631*) from Acheulean handaxes (*n = 10*) created by right- and left-handed subjects. Blinded handedness predictions for flakes were compared to their true handedness in order to assess each method's reliability. In order to test replicability, multiple observers classified a sample of flakes and inter-observer agreement was assessed. None of the methods were better than chance in predictive accuracy, and there were significant issues with inter-observer agreement. This study suggests that identifying knapper handedness in lithic debitage is extremely difficult, but also that some existing methodological issues may have simple solutions; suggestions for future research on this topic are provided.*

KEYWORDS: *Hominid handedness, Right shift, Lithic analysis, Experimental archaeology, Hominid cognitive evolution*

INTRODUCTION

Unique traits of *Homo sapiens*, such as bipedal locomotion, large and complex brains, culture, and language are typically cited as what make us *human*. One often-ignored aspect of human uniqueness, however, is handedness, although recent research has begun to highlight its evolutionary significance. Between 85 and 90 per cent of living humans are right-handed, and extreme, population-wide right-hand dominance is likely an important component of human singularity (Annett 1985; Corballis 1983, 1991). Over 150 years of research in the fields of anthropology, psychology, neuroscience, and others, have shown that manual motor function is linked to the neural areas responsible for human language, making handedness a viable proxy for understanding aspects of brain lateralization, technological advancement, and language acquisition in hominids (see Bruner 2015; Corballis 2003; Davidson 2010; Hewes 1993; Rilling 2008). Although we understand much about the link between handedness, cerebral lateralization, and language in living humans, little is known about the mechanisms involved in the evolution of these traits (Ruck 2014b).

Handedness has been extensively studied in humans, non-human primates, and even fossil remains, and one major goal of studying handedness in these contexts is to determine if right-hand predominance is unique to the hominid lineage, or if it has origins earlier in primate evolution, as do many other traits (Balzeau et al. 2011; Cantalupo and Hopkins 2008; Hopkins and Rilling 2000; Matsuzawa 2001; Sherwood et al. 2003). Intense debates exist in the literature on primate handedness, but the overall consensus is that manual motor lateralization is present in many non-human primate species, but there is no evidence of right-hand predominance in wild primate populations. Instead, we see that in many populations where most individuals show hand preference across all tasks, the preference is equally distributed between the right- and left-hands (Byrne 2005; Fagot and Vauclair 1991; Hopkins et al. 2007, 2011; McGrew and Marchant 1996, 1997). This suggests that while handedness may not be entirely unique to hominids, the selective

DOI: 10.1179/2051618515Y.0000000009

pressures specifically for right-hand dominance occurred sometime after the hominid lineage split, and we should therefore be able to track a *rightward shift* in hand preference in the fossil record (see Annett 1985, 1998; Cashmore 2009).

Many studies on skeletal bilateral asymmetry, which can preserve hand preference based on life-long differential biomechanical loading, have shown that right-handedness was quite common by the Upper Paleolithic (Lazenby 2002; Liguria 1997; Shaw 2011; Steele 2000; Steele and Uomini 2005, 2009; Trinkaus et al. 1994; Ubelaker and Zarenko 2012; Vandermeersch and Trinkaus 1995). However, in order to assess this asymmetry, researchers use paired long bones from a single individual, and as we move further back in time these become increasingly rare. Other data, such as directional skew in dental wear, cutmarks on bones from hominid butchery sites, Upper Paleolithic cave art, and asymmetries in hominid endocasts, also indicate that right-hand predominance was present in Neanderthals and perhaps other later *Homo* species (Bax and Ungar 1999; Bermúdez de Castro et al. 1988; Bromage and Boyde 1984; Faurie and Raymond 2004; Fox and Frayer 1997; Holloway 2008; Lozano et al. 2009; Pickering and Hensley-Marschand 2008; Uomini 2011; Volpato et al. 2012). Many believe that the prevalence of data for right-hand predominance in the Upper Paleolithic indicates that we can find even earlier evidence of the right shift in hominids.

Extending current knowledge on human handedness and brain lateralization backward into the fossil record has been extremely difficult due to an inherent issue in paleoanthropology: preservation. We do not have a solid timeline for spoken language acquisition in hominids, for example, because words do not fossilize. This is the case for many of the characteristics that distinguish *Homo sapiens* from other primates, such as cognitive complexity and cultural innovation. Despite decades of fieldwork, paleoanthropologists have relatively little primary evidence in the form of fossil data, and thus must develop methodologies that allow them to glean as much information as possible from small sample sizes (Cashmore et al. 2008; Holloway 2008; Schick and Toth 2006; Toth and Schick 1986; Wynn 2002).

Preservation will always be an issue for paleoanthropologists, but some aspects of hominid evolution preserve relatively well, including stone tools. Lithic technology dates from around 2.5 million years ago, and is directly associated with several hominid species, including all later species of the genus *Homo* (Ambrose 2001; Bordes 1968; Toth and Schick 1986). While much is currently known about the manufacture and use of Paleolithic stone tools, paleoanthropologists are just beginning to develop approaches relevant for studying handedness and brain lateralization via stone tools.

Flintknapping experiments have historically been useful for determining manufacture techniques and tool functions, but more recent studies focus on understanding the cognitive mechanisms that underlie stone-tool manufacture (see Bruner 2015; de Beaune et al. 2009; Coolidge and Wynn 2009; Gibson and Ingold 1993; Noble and Davidson 1996; Nowell and Davidson 2010; Roux and Bril 2005). This cognitive shift in experimental archaeology led to several studies on handedness in the hominid fossil record in the 1980s and 1990s, with a recent revival in interest in the twenty-first century (for a review, see Uomini 2001, 2006, 2009). Studies on the evolution of handedness are limited in number, and only a few of them reflect experimental approaches to inferring handedness from lithic evidence (Bargalló and Mosquera 2013; Patterson and Sollberger 1986; Pobiner 1999; Rugg and Mullane 2001; Toth 1985; Uomini 2001, 2006). These publications are speculative in nature and often debated, but they represent innovative approaches in paleoanthropology and form the basis of this study.

PREVIOUS WORK

Perhaps the earliest speculation about determining handedness from lithic materials was by S. A. Semenov, in his experimental studies on various Paleolithic tools. Semenov (1964) was particularly interested in the biomechanics of stone-tool production, and thus had an inherent interest in handedness. There are several inferences to the right-handedness of Paleolithic flintknappers throughout his work (Semenov 1964). Other than Semenov's work, lithic-based approaches to hominid handedness were non-existent until N. Toth, an American paleoanthropologist and flintknapper, published "Archaeological Evidence for Preferential Right-handedness in the Lower and Middle Pleistocene, and Its Possible Implications" in 1985.

In the article, Toth (1985) argued that the orientation of cortex on successively removed lithic flakes can be used to infer handedness from lithic assemblages. Toth applied his method, which was informed by experimental re-creation of simple core-scrapers, to assemblages from Koobi Fora, Kenya, and found similar right-to-left flake ratios. He concluded that this was initial evidence of high rates of right-handed knappers in early hominid populations, including *Homo habilis* and *Homo erectus* (Toth 1985).

This publication inspired much excitement about the possibility of early hominids being more behaviorally similar to modern humans than previously thought; however, debates regarding the integrity of Toth's (1985) method ensued. Two critiques of this study note that Toth's main assumption (of clockwise rotation of the core in the left hand of right-handed knappers and counterclockwise core rotation for opposite knappers) is not necessarily true of all knappers (Patterson and Sollberger 1986). Several attempts have been made to reassess Toth's technique, all of which have showed variable success in designating handedness to single and multiple knappers (see Pobiner 1999; Uomini 2001, 2006).

Outside of Toth's study (and the reviews it triggered), no attempts were made to identify handedness in lithic materials for decades. However, in a (1986) monograph on La Cotte de Saint Brelade, a Paleolithic site in Jersey, J. M. Cornford published evidence of handedness in lithic resharpening techniques (Cornford 1986). There is extensive literature on differential retouch of lithic materials, particularly of hafted tools, and other evidence, such as "prehensility" in Upper Paleolithic and Mesolithic tools, has reinforced these investigations of handedness in lithic remains (see Peresani and Miolo 2012, for an example). However, much like skeletal studies, the assemblages that could be studied regarding prehensility, retouch, and hafting are often later in time, when established evidence of right-handedness is already strongly evidenced (Frayer et al., 2012).

In 2001, G. Rugg and M. Mullane published a new method for inferring hominid handedness via stone tools, based on the skew of the cone of percussion on lithic flakes. They analyzed flakes made by both novice and expert subjects, and claimed that an overall rightward skew in the cone of percussion for flakes of a single knapping event successfully predict a right-handed knapper, and the reverse for a left-handed one. Rugg and Mullane correctly identified 75 per cent of flakes that had directional skew, but a majority of the flakes produced (roughly 85 per cent) showed no skew in the cone of percussion, so the low number of assignable flakes is discouraging. Despite these issues, Rugg and Mullane conclude that the two methods (theirs and Toth's) should be used in conjunction with each other, and that the chances of both methods being incorrect would be much smaller, leading to better accuracy overall (Rugg and Mullane 2001).

N. Uomini's unpublished studies (M.Sc., 2001, and Ph.D., 2006) on handedness and Paleolithic materials reflect the most significant reviews of both theory and practice in assessing hominid handedness in lithics. In her 2001 thesis, Uomini attempted to reproduce Toth's (1985) and Rugg and Mullane's (2001) studies. She noted that "both [...] method[s] were quite easy to learn and apply to large numbers of archaeological specimens. However, the reliability of the methods proved disastrous" (Uomini 2001:69). Uomini applied each method to 664 flakes from British Lower Paleolithic sites, Swanscombe and Purfleet, and none of her resultant ratios paralleled those from other publications (Uomini 2001). In fact, she concluded that the flakes from both sites seemed to be distributed to left- and right-handedness randomly, implying that no right shift was present in hominids at the sites (dated between 400 and 200 kya), or that the methods developed by Toth and Rugg and Mullane were ineffective at determining handedness (Uomini 2001).

Despite this, Uomini continued research on lithic indicators of handedness, and her dissertation (2006) was on inferring handedness from two more British Lower Paleolithic sites—Boxgrove and High Lodge. Uomini also used comparative data from 12 living knappers in this study, and conducted an extensive analysis of the flintknapping processes for each individual via interviews and video recordings. All subjects had experience knapping, although it was variable, and she simply asked knappers to produce "[...] what they felt capable of [...]" (Uomini 2006:78). Once again, Uomini had variable success using Toth's cortex model, and she found an even distribution of right- and left-skewed cones for most knappers in the experimental assemblage.

In an attempt to explain the weaknesses of each method, Uomini assessed the video data from each knapper, and found that flintknapping is highly stylistic and individualized. Knappers each had their own techniques for manipulating cores, which were often incompatible with the assumptions of both Toth's and Rugg and Mullane's models. Because the assumptions and validity of each model were refuted by the experimental data, Uomini did not attempt either method on the Lower Paleolithic assemblages, and instead focused on assessing Cornford's (1986) method using tranchet flakes. Despite her unsuccessful attempts in these studies, Uomini still advocates experimental approaches to studying hominid handedness, and suggests future research involving scanning electron microscopy and other new approaches in favor of replicating previous studies (Uomini 2006).

The final publication relevant to this study is "Can hand laterality be identified through lithic technology?" by Bargalló and Mosquera (2013). Bargalló and Mosquera assessed Toth's (1985) and Rugg and Mullane's (2001) methodologies, with some modifications, but they also introduced five new potential indicators of handedness in lithic remains: hackles and eraillure scars, structure and inclination of the striking platform, location of the impact point on platform, and locations of fractures on broken flakes (Bargalló and Mosquera 2013:8). In their study, Bargalló and Mosquera (2013) focused solely on experimental verification of knapper handedness. However, due to lack of expert knappers, they had to use novices, and they only required subjects to produce flakes (Bargalló and Mosquera 2013).

According to Bargalló and Mosquera (2013), Toth's (1985) and Rugg and Mullane's (2001) methods were unsuccessful in indicating knapper handedness, both independently and in conjunction with each other. A correspondence analysis on their own features revealed that no single trait on lithic flakes is accurate in predicting handedness, but that a combination of traits as a whole can predict knapper handedness relatively well. They note that, in general, left-skewed characteristics indicate a left-handed knapper, whereas right-skewed characteristics indicate a right-handed one. Several flakes analyzed showed a mixture of these features, however, and Bargalló and Mosquera state that "single flakes cannot be ascribed with certainty to a right- or left-handed knapper" as well as entire assemblages can (Bargalló and Mosquera 2013:24). Bargalló and Mosquera conclude by urging further studies, particularly ones with expert knappers.

METHODS

The purpose of this study was to test three previously established methods for determining handedness in lithic materials: Toth's (1985) cortex model, Rugg and Mullane's (2001) cone of percussion model, and Bargalló and Mosquera's (2013) technical features model. Under the assumptions of these works, right- and left-handers create predictable characteristics in the lithic debitage they produce, based on differential conchoidal fracture (see Cotterell and Kamminga 1990). However, based on the mixed results from these studies, it is clear that more work is needed to confirm or falsify each method's efficacy and replicability.

All lithic materials analyzed in this study were produced by volunteer modern-day expert flintknappers, as part of the first author's master's thesis research (Ruck 2014a). This study was approved by the Institutional Review Board at Florida Atlantic University (IRB ID No. 518291-1), and all subjects gave written consent prior to the study. Originally, two right-handed and two left-handed expert subjects—classified as those who have been regularly flintknapping for at least 5 years—were asked to create Acheulean handaxes in separate knapping sessions. In order to avoid ascribing meaning to individualized characteristics/variation, more than one subject of each hand preference was needed. Acheulean handaxes were chosen because they are restricted to the Lower Paleolithic period and are relatively easy to make, but they reflect a major technological transition in hominid evolution and have more complex manufacture methods than those involved in Oldowan tool manufacture (Hopkinson and White 2005; Schick and Clark 2003; Stout and Semaw 2006; Stout et al. 2006, 2009). There is also a general consensus that the shift from the Oldowan to the Acheulean coincides with significant cognitive expansion in hominid species, which has been supported by multiple studies on stone-tool production (Geribàs et al. 2010; McPherron 2000; Moore 2011; Stout 2005; Stout and Chaminade 2012; Stout et al. 2000, 2008; Uomini and Meyer 2013).

Materials were also provided to each subject in order to maintain control over external sources of error (i.e. differing conchoidal fracture patterns, hardness of rocks used, amount of initial reduction required, etc.). Although consistency is difficult when dealing with lithic raw materials, the same source materials were used in an attempt to establish as much experimental control as possible. Each subject received Edwards Plateau chert nodules and one hammerstone, all from a single source. Subjects created their handaxes while sitting, using direct percussion. They were directed to create handaxes between 9 and 12 cm in maximum dimension, instead of producing a set number of flakes or completing a set number of blows, in order to maintain a realistic mixed meta-assemblage. All materials produced in the creation of the tools (including the finished product, flakes, hammerstones, shatter, etc.) were collected by the subjects and returned to Florida Atlantic University for analysis. Once received, the products of each knapping session were kept separated by handaxe and by subject in order to facilitate cataloging.

The original intent was to get 12 handaxes made by four subjects, but one right-handed subject was unable to complete one handaxe due to a low-quality nodule, and one left-handed subject was unable to complete any handaxes for unknown reasons. Because of these issues, two additional left-handed subjects were opportunistically recruited at a knap-in in northern Florida, and each made one Acheulean handaxe, also using Edwards Plateau chert. One knapper made the handaxe on-site using only a copper billet, and the other later sent the handaxe and debitage to Florida Atlantic University like the originally recruited subjects. The same cataloging and analytical methods were used on these materials.

In sum, 10 handaxe "assemblages" were made for this thesis, five by right-handers and five by left-handers. Upon arrival at Florida Atlantic University, all flake debitage for each handaxe was sorted and labeled by two undergraduate anthropology students. Each flake over 2 cm in minimum dimension was labeled with a random number between 1 and 1000. Along with the flake number, each flake's associated handaxe, the subject it came from, and finally, the true handedness of the flake (i.e., the handedness of the subject) were also recorded. After each handaxe assemblage was coded, all flakes were mixed together in a single "meta-assemblage." The first author then conducted a blind analysis on the entire coded meta-assemblage using Toth's (1985), Rugg and Mullane's (2001), and Bargalló and Mosquera's (2013) techniques. Data recording and analysis were facilitated using a Microsoft Access database, and then exported to Microsoft Excel and eventually SPSS for analysis.

The first handedness analysis conducted was based on the cortex model introduced by Toth in 1985, so only cortical flakes were analyzed. Using judgment by eye, cortex was designated as right-oriented or left-oriented based on predominance of cortical surface area. The next analysis replicated Rugg and Mullane's cone of percussion method from 2001. Only flakes with identifiable platforms and observable cones of percussion were used for this method. Thus, cones and ridges were both simply categorized by inspection as left-skewed, right-skewed, or centered. The last method was derived from Bargalló and Mosquera's (2013) publication, which is also a review of Toth's (1985) and Rugg and Mullane's (2001) methodologies. This analysis was the central component of this study, and was completed on all cataloged flakes. Data were recorded using a Microsoft Access database, which included the following 15 fields, derived from Bargalló and Mosquera's original paper: cone of percussion, hackles, ripples, extraction axis, ridge angle, platform type, platform inclination, impact point, cortex, fracture locations, and eraillure scar locations.

Bargalló and Mosquera (2013) found that some of these characteristics did not show clear separation between right-handed and left-handed knappers, so an informal pilot study on flakes from two novice, female subjects (one right- and one left-handed) was used to supplement Bargalló and Mosquera's handedness classifications. Based on these data, right-handed designations were usually made for flakes with a right-extraction axis (includes rightward cone of percussion, right impact point, and fractures D and G) and left-oriented eraillure scars, while left-handed designations were usually made for a left-extraction axis (with left-skewed cones and impact points, and fractures at E and F). As Bargalló and Mosquera found, however, many flakes often exhibited a mosaic of these traits (see Figure 1, as well as Bargalló and Mosquera (2013:8) for an additional description of these methods).

For every flake, five independent handedness deductions were made, one for each of the following groups of characteristics: cone of percussion, platform, cortex, fracture locations, and eraillure scar locations. Based on these five predictions, an overall handedness inference was made for every flake: "right," "left," or an "indeterminate" option for ambiguous flakes or cases where some observations indicated right-handedness while others indicated left-handedness on the same flake.

After all flakes were analyzed by the first author, the original coding catalog was used to regroup flakes by handaxe and finally by knapper, reconstructing the initial assemblages. In total, 631

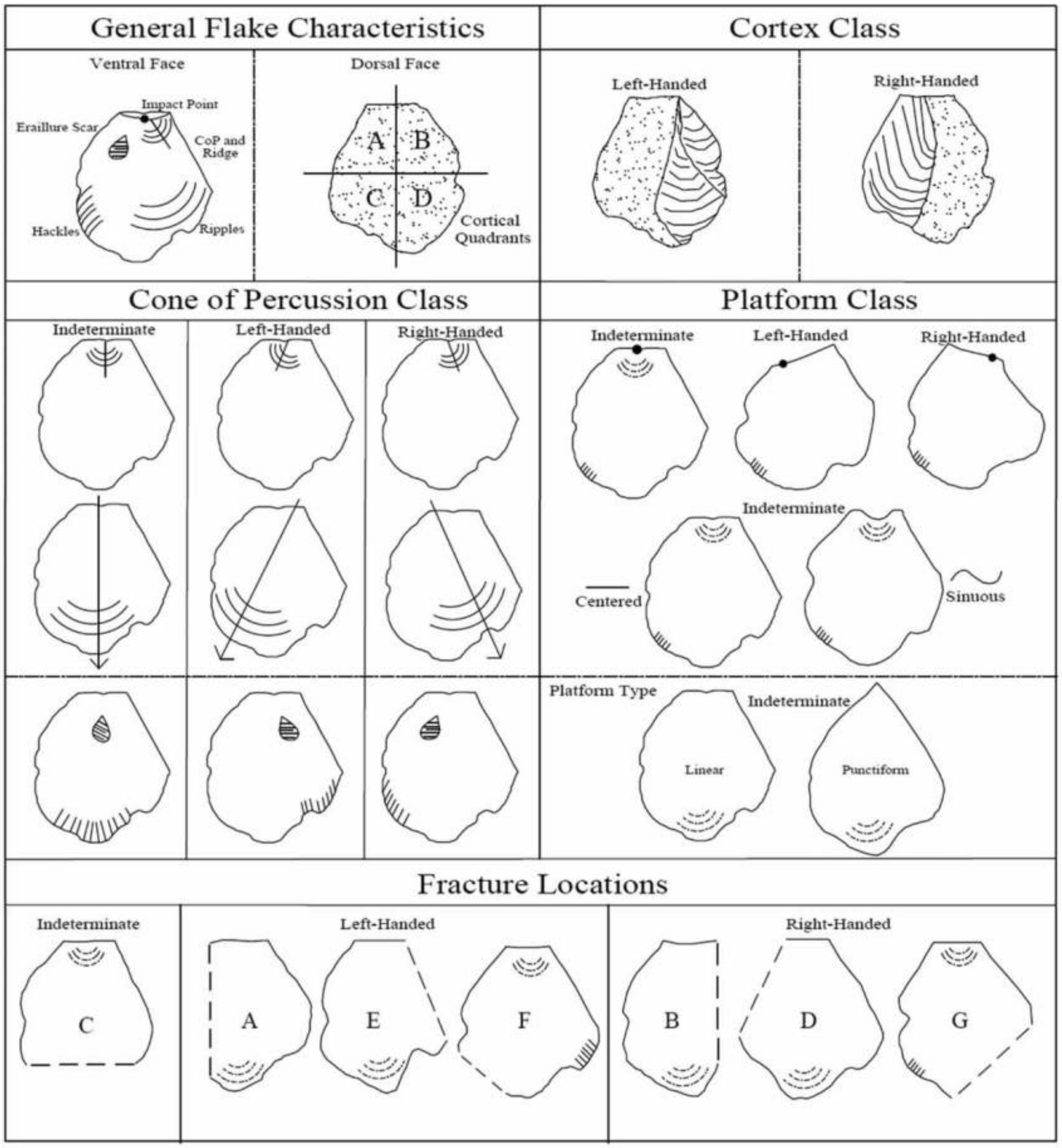

FIGURE 1. Description of flake characteristics and their associated handedness inferences, modified from Bargalló and Mosquera (2013).

flakes were analyzed as part of this study. It is important to note that only 242 (38.4 per cent) flakes came from right-handed knappers, while 389 (61.6 per cent) came from left-handers, simply because the latter produced more flakes per handaxe on average. Table 1 shows the number of flakes included in this study by subject, handedness, and handaxe.

These data were collected as part of a master's thesis, and thus, the sample size ($n = 631$) was deemed appropriate for the original context of evaluating the three reviewed methods for determining handedness in lithic debitage. However, some issues with the sample must be addressed, most notably, the differential contribution from subjects (consider that two left-handed subjects made only one handaxe), and the low number of flakes for some handaxes (such as B and F, both from a single knapper). Low n's become particularly problematic for some methods, such as Rugg and Mullane's (2001) cone of percussion approach, because a majority of the flakes in the general sample do not exhibit the characteristics necessary for analysis. Still, we believe that the knowledge gained from our approach outweighs the low statistical power within the sample.

Many weaknesses present in the previous studies can also be ascribed to observational biases (see

TABLE 1 META-ASSEMBLAGE BREAKDOWN BY HANDAXE AND KNAPPER

Handaxe	Knapper	Knapper handedness	Number of flakes	% of meta-assemblage
A	SH	Right	62	9.8
C	SH	Right	89	14.1
B	MC	Right	16	2.5
E	MC	Right	64	10.1
F	MC	Right	11	1.7
D	EM	Left	100	15.8
H	EM	Left	57	9.0
I	EM	Left	92	14.6
G	OS	Left	62	9.8
J	RC	Left	78	12.4

Rugg and Mullane 2001:254–255, for a discussion on observer objectivity). Therefore, after the first author completed all analyses, one anthropology graduate student and one anthropology faculty member (the third author)—both of whom are skilled in lithic analysis—also conducted a blind analysis on a stratified random sample of complete flakes, using only Bargalló and Mosquera's methods. After the three independent analyses were completed, cross-comparisons were made for each observer using Fleiss' Kappa and weighted Cohen's Kappa (Ben-David 2008; Fleiss et al. 1969; Nichols et al. 2011; Warrens 2011). The motivation for this aspect of the methodology is that many of the characteristics described by Bargalló and Mosquera (2013) appear ambiguous or subjective in nature, and no previous attempts to address inter-observer error in determining handedness from lithic debitage have yet been made.

In sum, this study was an attempt to answer several main questions regarding the existing literature on determining handedness in lithic debitage:

(1) Do Toth's (1985), Rugg and Mullane's (2001), and Bargalló and Mosquera's (2013) methods reliably indicate knapper handedness in a mixed assemblage of single lithic flakes, when employed by a single assessor?
(2) Do multiple assessors classify flakes in the same way, or is observer bias a problem within Bargalló and Mosquera's (2013) approach?
(3) Is handedness obscured by other factors, such as individual knapping style? Can these factors be assessed, or even controlled for, in future works?

RESULTS

The following handedness inferences for each handaxe were made based on whether a majority of the flakes for that handaxe were designated right- or left-skewed. For Toth's (1985) method, this relied upon the distribution of cortex on cortical flakes, where predominance of cortex was classified as left or right. Table 2 shows the frequencies of left- and right-classified cortical flakes by handaxe (for "observer A," or the first author); 159 flakes were not included in this analysis because they were fragments or shatter that could not be oriented properly, and 251 flakes were either fully cortical or non-cortical, and thus did not aid in classifying handedness. In total, only about 35 per cent of the flakes collected were viable for analysis using this method.

Using this method, six handaxe assemblages were correctly predicted, while four were not. Note that the binomial probability of randomly obtaining exactly four wrong and six right answers in 10 attempts is $P = 0.205$, and the cumulative binomial probability of getting six *or more* correct answers at random is $P \approx 0.377$. Thus, obtaining six correct inferences could easily be pure luck. Of the assemblages that were incorrectly predicted, two had relatively equal percentages of right- and left-cortical flakes, but all other assemblages had a clear predominance of either right- or left-cortex. It must be addressed that inferring handedness from simple predominance within each handaxe assemblage is likely too simplistic, especially in cases where so few flakes were viable for analysis (as in handaxe F, $n = 5$).

In Toth's original publication (1985), he interpreted two assemblages from Lower Paleolithic

TABLE 2 SUMMARY OF THE TOTH CORTEX-BASED METHOD

Handaxe	Knapper handedness	*n* of flakes		Percent		Handedness inference	Correct?
		Right	Left	Right	Left		
A	Right	5	12	29.4	70.6	Left	No
C	Right	13	10	56.5	43.5	Right	Yes
B	Right	5	3	62.5	37.5	Right	Yes
E	Right	4	9	30.7	69.3	Left	No
F	Right	4	1	80.0	20.0	Right	Yes
D	Left	6	12	33.4	66.6	Left	Yes
H	Left	6	4	60.0	40.0	Right	No
I	Left	8	10	44.5	55.5	Left	Yes
G	Left	7	5	58.4	41.6	Right	No
J	Left	7	11	38.8	61.2	Left	Yes

sites as having right-hand dominant knappers based on his own assemblage of flakes, which roughly showed a 55:45 right-to-left ratio of cortical skew. However, in our recreation of his methods, assemblages with even stronger directionally skewed ratios (e.g., handaxes E and G with an approximate 60:40 right-to-left distribution, or handaxe C with a 70:30 left-to-right skew) showed handedness *opposite* to the ratio, suggesting that the ratios are also unreliable predictors of handedness. Furthermore, one might predict that if the methods were working correctly, then assemblages with large numbers of flakes, that is, larger samples, would have been more likely to have been correctly identified. The opposite was actually the case: the two smallest assemblages, B and F, were correctly identified. The mean number of flakes in the correctly ($\bar{x} = 15$) and incorrectly ($\bar{x} = 15$) identified assemblages were similar. A Mann–Whitney U test does not find a significant difference between the numbers of flakes ($P = 0.45$). It is possible that the method works, but only weakly, so a much larger number of experiments are needed to discern the effect statistically. Coupled with the other critiques of Toth's method, this reassessment the cortex method suggests that is inaccurate at predicting knapper handedness, and considering that it could only be applied to a minority of the flakes, it is not sufficient as a stand-alone approach for assessing handedness in lithic debitage.

For Rugg and Mullane's (2001) method, the assessment was based only on right- or left-skew of the cone of percussion, as the first author also found the protractor method to be unemployable (see Rugg and Mullane 2001:253). Only 49 of the 631 (7.7 per cent) flakes analyzed had visible ridges, and many of these, as Rugg and Mullane stated, did not have platforms that facilitated the protractor method. Additionally, 255 flakes were fragments or shatter where the skew of the cone could not be identified, and 308 had centered cones of percussion. In sum, only about 11 per cent of the flakes analyzed indicated knapper handedness via the cone of percussion. Table 3 shows the frequencies of left- and right-skewed flakes by handaxe using Rugg and Mullane's (2001) method.

Using this method, only four handaxe assemblages were predicted correctly, and the number of flakes that showed clear directional skew in the cone of percussion was far too small for this method to be beneficial on its own. The binomial probability of randomly obtaining exactly six wrong and four right answers in 10 attempts is $P = 0.205$, while the cumulative binomial probability of getting four *or more* correct answers at random is $P \approx 0.828$. So, obtaining four correct inferences could be attributable to luck, and not even very good luck. Once again, it must be mentioned that our sample size, particularly for this assessment of Rugg and Mullane's method, is very small, and, as mentioned in our review of Toth's (1985) method, predominance is a very simple way of inferring handedness from this characteristic. The authors noted that the chance of both Toth's (1985) method and their own being incorrect when combined should be very low, but in many cases, the combination of these methods leads to contradictory or still-incorrect handedness assumptions as well (see Tables 2 and 3). These results suggest that both methods are inappropriate for use on single flakes, and perhaps on well-associated debitage as well.

TABLE 3 SUMMARY OF THE RUGG AND MULLANE CONE OF PERCUSSION METHOD

Handaxe	Knapper Handedness	Flake frequency		Percent		Handedness Inference	Correct?
		Right	Left	Right	Left		
A	Right	0	3	0.0	100.0	Left	No
C	Right	7	1	87.5	12.5	Right	Yes
B	Right	2	2	50.0	50.0	Indeterminate	No
E	Right	2	6	25.0	75.0	Left	No
F	Right	0	1	0.0	100.0	Left	No
D	Left	5	8	38.4	61.6	Left	Yes
H	Left	3	1	75.0	25.0	Right	No
I	Left	10	5	66.6	33.4	Right	No
G	Left	1	5	16.6	83.4	Left	Yes
J	Left	2	4	33.4	66.6	Left	Yes

Finally, the first author assessed the methods introduced in Bargalló and Mosquera's (2013) publication, on all 631 flakes. In general, a majority of flakes had many missing features, or a mixture of right- and left-skewed traits, resulting in a strong predominance of "Indeterminate" handedness classifications, labeled as IND (see Table 4: center columns). In fact, indeterminate classifications represented over 50 per cent of the classified flakes for seven out of the 10 handaxes. The average number of flakes classified correctly as right- or left-handed per handaxe ($\bar{x} = 13$) does not differ from those predicted incorrectly ($\bar{x} = 12$). Much like our assessment of Toth's method, a Mann–Whitney *U* test does not find a significant difference between the numbers of flakes ($P = 0.67$). Furthermore, no major differences in predictive success exist between handaxe, handedness, or knapper. It is important to note that the predictive correctness of each subset (for example, using only the cone of percussion subset, or only the fracture locations) was also discouraging, and no characteristic seemed to inform predictions better on its own. Using the five subsets in a combined overall handedness inference had little impact on the accuracy of the judgments, but in some cases, it did reduce the number of indeterminate flakes.

As stated by Bargalló and Mosquera (2013), many lithic flakes showed a mosaic of traits, making flake-by-flake interpretations very difficult. They suggest that their method would be most effective on well-preserved knapping scatters, where all flakes can be analyzed with reference to each other. As mentioned before, however, preserved paleo-archaeological scatters are extremely rare in comparison to mixed assemblages, so we wanted to test all methods on single flakes in this study.

TABLE 4 SUMMARY OF THE BARGALLÓ AND MOSQUERA TECHNICAL FEATURES METHOD

Handaxe	Knapper handedness	*n* flakes	*n* correct	% Correct	*n* IND	% IND	*n* incorrect	% Incorrect
A	Right	62	8	12.9	39	62.9	15	24.2
C	Right	89	17	19.1	60	67.4	12	13.5
B	Right	16	8	50.0	2	12.5	6	37.5
E	Right	64	12	18.8	32	50.0	20	31.2
F	Right	11	3	27.3	5	45.5	3	27.2
D	Left	100	21	21.0	59	59.0	20	20.0
H	Left	57	9	15.8	41	71.9	7	12.3
I	Left	92	21	22.8	48	52.2	23	25.0
G	Left	62	10	16.1	48	77.4	4	6.5
J	Left	78	25	32.2	40	51.2	13	16.6

As suggested by the predictive correctness of each method, analyzing lone flakes by simple judgment using these techniques is an unreliable method for determining handedness. However, these assessments reflect only a simple recreation of each methodology, and coupled with the inherent issues with sample size, it is possible that other, more statistically robust approaches may work better. Each method's inaccuracy in informing handedness inferences could be due to improper judgment of flakes with mosaic traits, or it could also be due to poor handedness associations for some of the traits. For example, the association between platform inclination and handedness was relatively hard to make because it had shown no clear differentiation in either Bargalló and Mosquera's (2013) or our own pilot study, so we rarely knew how to interpret it.

INTER-OBSERVER COMPARISONS

As mentioned before, every flake characteristic introduced by Bargalló and Mosquera (2013) is categorical in nature, and each trait reviewed in this study is simply judged by eye; this introduces problems with replicability, particularly between observers. In order to test whether different observers classify flake characteristics in the same manner, two additional observers who specialize in lithic analysis assessed a stratified random sample of flakes. All three observers assessed 25 flakes in common, and an additional 25 flakes were assessed by only two observers. Thus, observer A (the first author) assessed all 631 flakes, and observers B and C each assessed a stratified random sample of 50 complete flakes (shatter and extremely fragmented, "difficult-to-classify" flakes were not sampled), 25 of which overlapped so all three assessors had classified them. Before testing inter-observer agreement, we assessed how well each observer identified knapper handedness for their sample of 50 flakes using Bargalló and Mosquera's (2013) methods (see Table 4 of the first author's assessment of Bargalló and Mosquera's methodology, for comparison). Tables 5 and 6 show how well each observer classified their flakes as right- or left-handed by handaxe.

Both observers show results much like the full analysis, with "Indeterminate" handedness inferences being quite common, and no significant differences between the number of correctly classified vs. incorrectly classified flakes, in many cases. As stated before, flakes often showed either a mixture of right- and left-associated characteristics, or a predominance of centered characteristics, often making handedness inferences rather difficult to make. In general, the additional observers found the classification scheme to be quite difficult to utilize, even though they were familiar with Bargalló and Mosquera's (2013) methodology. In fact, the target sample size for this study, which was originally 50 flakes overlapped between the three observers, and 100 flakes overlapped by each pair of observers, had to be reduced because of the time required for analysis. Overall, this confirms that knapper handedness is extremely hard to identify in single flakes using simple judgment with Bargalló and Mosquera's (2013) classification scheme. Much like the previously mentioned results, the small sample size should be increased in future studies, and future assessors should be familiar with all of the relevant literature on inferring handedness from lithic debitage.

TABLE 5 SUMMARY OF OBSERVER B'S HANDEDNESS INFERENCES

Handaxe	Knapper handedness	*n* flakes	*n* correct	% Correct	*n* IND	% IND	*n* incorrect	% Incorrect
A	Right	1	0	0.0	0	0.0	1	100.0
C	Right	2	1	50.0	0	0.0	1	50.0
B	Right	0	0	0.0	0	0.0	0	0.0
E	Right	8	1	12.5	3	37.5	4	50.0
F	Right	4	3	75.0	0	0.0	1	25.0
D	Left	6	2	33.3	2	33.3	2	33.4
H	Left	3	2	66.7	1	33.3	0	0.0
I	Left	7	4	57.1	2	28.6	1	14.3
G	Left	8	2	25.0	3	37.5	3	37.5
J	Left	10	5	50.0	2	20.0	3	30.0

TABLE 6 SUMMARY OF OBSERVER C'S HANDEDNESS INFERENCES

Handaxe	Knapper handedness	*n* flakes	*n* correct	% Correct	*n* IND	% IND	*n* incorrect	% Incorrect
A	Right	3	1	33.3	2	66.7	0	0.0
C	Right	2	0	0.0	2	100.0	0	0.0
B	Right	0	0	0.0	0	0.0	0	0.0
E	Right	7	1	14.3	5	71.4	1	14.3
F	Right	0	0	0.0	0	0.0	0	0.0
D	Left	6	3	50.0	2	33.3	1	16.7
H	Left	3	0	0.0	2	66.7	1	33.3
I	Left	8	2	25.0	3	37.5	3	37.5
G	Left	11	6	54.6	3	27.2	2	18.2
J	Left	10	4	40.0	4	40.0	2	20.0

For the 25 flakes analyzed by all three observers, a Fleiss' Kappa value was calculated to determine agreement, with values close to 0 indicating poor agreement and a value of 1 indicating full agreement between all three observers (Fleiss et al. 1969). Table 7 shows Fleiss values and standard error (SE) for each characteristic, as well as each subset handedness inference, and finally the overall handedness inference.

In general, Fleiss values suggest that inconsistency is indeed an issue when using Bargalló and Mosquera's (2013) methodology. Relatively few characteristics have values above 0.2, which reflects "partial agreement" between all three observers. Additionally, only four characteristics, all of which were fracture locations, have values above 0.5. It is important to note that the sample size for these tests was quite low at only 25 flakes out of 631 (less than 5 per cent), and some SE values are relatively large compared to their Fleiss values.

For the flakes that were only classified by two observers ($n = 50$), a weighted Cohen's Kappa test was used to assess inter-observer differences in classifications. Weighted Kappa values represent not only agreement between observers, but also the magnitude of disagreement (Nichols et al. 2011; Warrens 2011). Like Fleiss' values, 1 represents full agreement and values close to 0 represent poor agreement. Table 8 shows the weighted Cohen's Kappa values and associated SE for each characteristic and handedness inference.

As indicated by the weighted Cohen's Kappa values, no single observer seems to disagree with the other two more often across characteristics, and a sample size of 50 (versus 25) does not significantly impact the SE in most cases. Like the Fleiss tests, the weighted Cohen's Kappa values suggest that inter-observer subjectivity is a large issue with this methodology. Rugg and Mullane (2001) also addressed replicability by having the same observer look at each flake twice, with a few days in between judgments, so even for a

TABLE 7 FLEISS' KAPPA VALUES FOR EACH CHARACTERISTIC USING ALL THREE OBSERVERS, $N = 25$

Flake characteristic	Value*	SE
Cone of percussion	0.1232	0.0660
Hackles	0.1599	0.0659
Ripples	0.1587	0.0654
Extraction axis	0.0966	0.0793
Ridge	0.0113	0.0806
Eraillure center	0.0227	0.1155
Eraillure right	0.2424	0.1155
Eraillure left	0.3333	0.1155
Platform type	0.2110	0.0885
Platform inclination	0.0891	0.0671
Impact point	0.2526	0.0766
Cortex	0.3777	0.0498
A fracture	1.0000**	0.0000
B fracture	0.0135	0.1155
C fracture	0.5273**	0.1155
D fracture	0.4565	0.1155
E fracture	0.5273**	0.1155
F fracture	0.5210**	0.1231
G fracture	0.3608	0.1155
Overall handedness inference	0.2118	0.0836

*Values marked.

**Indicates at least moderate agreement between all three observers.

Table 8 Weighted cohen's kappa values for each characteristic using pairs of observers

Flake characteristic	Observer A–B*	SE	Observer A–C*	SE	Observer B–C*	SE
Cone of percussion	0.1483	0.1058	0.1250	0.1270	0.2778	0.1546
Hackles	0.3143	0.1057	0.0989	0.1445	0.1294	0.1721
Ripples	0.2161	0.0972	0.0935	0.0972	0.2629	0.1709
Extraction axis	0.0643	0.1114	0.0933	0.1113	0.2424	0.1751
Ridge	0.0329	0.0511	0.2444	0.1199	0.1596	0.0057
Eraillure center	0.2599	0.1396	0.1379	0.2296	0.3902	0.2170
Eraillure right	0.4595**	0.2592	0.1814	0.1953	0.1804	0.0174
Eraillure left	0.3961	0.2562	0.3386	0.2296	0.3067	0.2331
Platform type	0.2529	0.1959	0.3094	0.1628	0.2021	0.2840
Platform inclination	0.0721	0.1146	0.2770	0.1229	0.2097	0.0226
Impact point	0.2385	0.1278	0.2574	0.1216	0.1983	0.0741
Cortex	0.3222	0.1417	0.4086**	0.1219	0.5044**	0.1290
A fracture	1.0000**	0.0000	0.707**	0.0204	1.0000**	0.0000
B fracture	1.0000**	0.0000	0.4944**	0.0309	0.9798**	0.0012
C fracture	0.3450	0.2509	0.2105	0.2383	0.3595	0.3363
D fracture	0.1071	0.2525	0.3056	0.2021	0.3595	0.3363
E fracture	0.1935	0.2827	0.6575**	0.3390	0.6269**	0.2351
F fracture	0.1583	0.2253	0.3363	0.2542	0.7500**	0.1696
G fracture	0.2372	0.2302	0.2857	0.2504	0.4118**	0.2696
Overall handedness inference	0.1192	0.1200	0.1302	0.1248	0.2325	0.1748

*Values marked.

**Indicates at least moderate agreement between pairs of observers.

single observer, problems with replicability likely exist. These issues are abundantly clear when classifications from multiple observers are compared, and considering the general inefficacy of the predictions based on the full sample from a single observer, alternative approaches for data collection should be investigated.

EFFECTS OF KNAPPING STYLE ON LITHIC DEBITAGE

As suggested by Uomini's (2006) dissertation findings, it is likely that individualized knapping styles manifest in the flakes that knappers produce, ultimately leading to highly varied assemblages that are extremely difficult to evaluate. Informed by her work, we tested whether there were detectable differences in flake characteristics by *knapper*, instead of by handedness. For this analysis, a simple Pearson's chi-squared test was run on each of Bargalló and Mosquera's characteristics. Chi-squared values were calculated for each variable frequency by handedness first, which rarely resulted in significant *P*-values. However, when chi-squared values are calculated by knapper, almost every characteristic shows significant differences in frequencies at $a = 0.05$. Comparative chi-squared values are shown for each feature in Table 9.

The chi-squared tests suggest that knappers produce unique combinations of the technical features identified by Bargalló and Mosquera (2013), which opens up new avenues of research that are not necessarily related to identifying handedness. It has been repeatedly mentioned in previous studies that clusters of flakes are more viable for analysis than single flakes because they can be assessed within the context of each other. In many cases, however, clearly associated groups of flakes are rare. Although the major motivation of this study was to evaluate the existing methods for identifying handedness, a working method for associating multiple flakes to a single knapper may also be useful in assessing paleo-archaeological sites. Addressing the variability produced by individual knappers may, in fact, be a preliminary step for assessing handedness in debitage, although this question needs to be studied further in future works.

Table 9 Pearson's Chi-Squared tests by handedness and knapper

Flake characteristic	Handedness			Knapper		
	#χ^2	df	*P*-value*	#χ^2	df	*P*-value*
Cone of percussion	7.676	3	0.053	23.309	12	0.025**
Hackles	3.960	3	0.266	21.483	12	0.044**
Ripples	2.843	3	0.416	23.206	12	0.026**
Extraction axis	2.953	3	0.399	15.326	12	0.224
Ridge	36.942	3	0.000***	104.437	12	0.000***
Centered eraillure scar	0.270	1	0.603	4.524	4	0.340
Right eraillure scar	1.011	1	0.315	13.321	4	0.010***
Left eraillure scar	1.801	1	0.180	15.000	4	0.005***
Platform type	1.801	1	0.180	7.956	4	0.093
Platform inclination	17.714	3	0.001***	32.359	12	0.001***
Impact point	0.140	2	0.932	15.660	8	0.048**
Cortex location	9.476	15	0.851	62.104	60	0.401
A fracture	0.583	1	0.445	10.747	4	0.030**
B fracture	0.021	1	0.886	2.214	4	0.696
C fracture	4.962	1	0.026**	9.281	4	0.054
D fracture	0.887	1	0.346	5.128	4	0.274
E fracture	0.239	1	0.625	2.735	4	0.603
F fracture	1.180	1	0.277	10.848	4	0.028**
G fracture	1.842	1	0.175	5.785	4	0.216

*Values marked.
**Indicates significance at $a = 0.05$. Values marked.
***Indicates significance at $a = 0.01$.

Conclusions

Each of the methods reviewed as part of this study has various theoretical bases, motivations, and methodological approaches, but all of them have come to the same conclusions: definitive evidence of handedness is extremely difficult to find outside of studying living human beings, and lithic-based evidence of handedness is a complex issue. Our results reflect the limitations present in the studies herein reviewed, and we have also introduced some new topics that need addressing.

In the case of Toth's cortex method (1985), all analyses of cortex location show that it is unreliable at indicating knapper handedness. As mentioned earlier, Toth's original method (1985) assumes unidirectional rotation of the core in the knapper's non-dominant hand. This reduction method could perhaps be true in Oldowan manufacture, but decortication is an initial step in many biface manufacturing techniques, and it relies upon multi-directional rotation of cores, where left- and right-skew of cortex is perhaps simply unrelated to directional core rotation. In terms of inter-observer agreement, cortex showed relatively high agreement between raters, but this study suggests that cortex location should be abandoned within the context of studying handedness pending further work on decortication approaches for various tool types. Our replication of Rugg and Mullane's (2001) methodology also proved unsuccessful in terms of each observer's handedness judgments. The kappa values for ridge and cone of percussion were quite low, suggesting issues with replicability, and the repeated failure of the protractor method, which could eliminate some subjectivity, does not seem like an alternate approach. The cone of percussion method suggested by Rugg and Mullane (2001) should be used with caution, likely because its underlying assumptions regarding direction of percussion are also subject to knapping styles and factors other than handedness.

Our review of Bargalló and Mosquera's (2013) methodology supports their claim that the listed traits are inherent to lithic manufacture, and that they may be useful for detecting knapper characteristics. However, the high percentage of flakes with mosaic traits made handedness predictions extremely difficult to make, and some relatively "straightforward" flakes were still interpreted incorrectly. Poor handedness associations for some traits, as well as a lack of context for interpreting mixed characteristics, are perhaps the largest weaknesses of this approach. As suggested by the chi-square values, it seems that Bargalló and Mosquera's characteristics were produced in varying frequencies by the subjects in this study, and we strongly encourage future studies to reaffirm the relationship between manufacture style and debitage characteristics. Based on these data, perhaps identifying a cluster of flakes from a single knapper at a site could be a preliminary step for assessing the knapper's handedness.

Although flintknapping is a hobby for many, expert flintknappers interested in this type of research are rare, especially those who are left-handed (see Bargalló and Mosquera 2013). Additionally, studies that infer handedness from material culture have such low signal-to-noise ratios that they seem impractical to many, and the low number of assignable flakes produced from a single knapping event are discouraging (Patterson and Sollberger 1986; Pobiner 1999; Uomini 2001). Combining these inherent limitations with the complexity of lithic analysis and the breadth of Paleolithic assemblages, it is no surprise that there have only been a handful of studies ever conducted. Some basic considerations, specifically introduced within this study, follow: first is that evidently, the theoretical premises for flake characteristics within the three reviewed methods do not correlate very well with our meta-assemblage, and it is clear that personalized knapping styles and other factors lead to highly variable data. The influence of knapping style on technical characteristics in flakes was heavily discussed by Uomini (2006), but it is relatively unaddressed in the wider literature. Based on our study, future works for detecting individual knappers in an assemblage may be extremely fruitful.

Perhaps another reason for the variability in the present data is a lack of experimental control. In terms of this study, it is unclear whether the frequencies of some features (e.g. fractures, ripples, hackles) may be related to the material of the nodule vs. the knapper vs. handedness itself. Some recent studies have suggested using alternate materials, such as silicate-based bricks, in knapping experiments as additional experimental control (see Geribàs et al. 2010; Khreisheh et al. 2013). The use of uniform porcelain blanks could eliminate, or at least reduce, the effect of material inconsistency on flake characteristics, which would be extremely helpful in studies on differential fracture mechanics of right- and left-handed knappers, especially considering that these materials still preserve all aspects of worked stone (including traits of conchoidal fracture from manufacture, and even use-wear (Khreisheh et al. 2013).

Another source of unaddressed variation in the present meta-assemblage likely stems from the generally relaxed manufacture guidelines we gave to the participants, which left styles largely in the control of the subjects themselves. One motivation of this study was to produce results that could be applied to actual assemblages, so we did not impose strict guidelines that would result in a meta-assemblage unlike those we find at typical sties. Overall, however, this had large unintended effects: a balanced 50–50 distribution was not achieved, and larger differences by subject weakened our analyses as well. For example, the only left-handed knapper to make all three handaxes produced 249 of the flakes analyzed in this study (almost 40 per cent of the meta-assemblage), whereas the only right-handed knapper that made all three handaxes only produced 91 flakes (roughly 14 per cent) above 2 × 2 cm (see Table 1). Additional controls on the creation and collection of materials may help future researchers obtain a better data set, which may lead to better associations between flake characteristics and handedness, through a reduction of external variability.

Furthermore, we believe that many of the problems encountered in work of this nature are consequences of the nominal nature of the reviewed classification systems. Cataloging features as nominal data is certainly the most straightforward approach, but it severely limits the avenues for statistical analysis and evaluation. Issues with the nominal classification scheme are particularly apparent in the inter-observer differences for Bargalló and Mosquera's methodology. Rugg and Mullane (2001) attempted to address this by introducing the protractor method and assessing each flake twice, but they were unsuccessful in these endeavors (Rugg and Mullane 2001). As indicated by the Fleiss' and weighted Cohen's Kappa values for each characteristic (see Tables 7 and 8), multiple raters do not typically characterize a single flake as having the same characteristics. Considering the nature of the classification system, where all assessments are made by simple eye judgment and the options within each independent variable (center, left, right, or none, in most cases) are hard to distinguish for many flakes, this result is not surprising. Future replications of Bargalló and Mosquera's (2013) methodology should, at the very least, have three observers classify all flakes (instead of just a sample), and in the case of a one-way disagreement, the modal classification should be used for that flake. In a case where all three observers disagree on a characteristic, it should not be used in analysis, although this was relatively rare in our sample.

Another possible solution to the issue of unreliability is much like Uomini's final recommendation in her dissertation, although she was not proposing it within the same context. We recommend that future studies use a more rigorous morphotechnical measurement system, such as 3D scanning, instead of simple observation for characteristics on flakes. In general, nominal data are hard to analyze, and outside of analyzing frequencies, little has been done to test the existing data collection schemes. Spatial data are relatively complex, but they can be manipulated and statistically evaluated in a multitude of ways, which may lead to better associations between traits and handedness or knapper. Second, 3D scanning could significantly reduce the effect of observer bias in data collection, although some subjectivity may still be inherent in rendering scans. An approach using 3D scanning, however, would require a major modification of the existing classification schemes, if not an entirely new system. Due to the involved nature of scanning, this approach should first be used on a small sample of flakes, perhaps with the simple goal of identifying how well Bargalló and Mosquera's (2013) characteristics are exhibited in scans, and checking if these scans improve predictive correctness and multi-rater agreement. If improvements are clear, a more involved scanning project could be viable.

Despite the issues encountered, we believe that lithic analysis within the context of handedness could provide exceptional insights into hominid evolution, if these methods continue to be improved upon. Based on this study, it is apparent that future works need a much larger, and more balanced sample of flakes, in order to discern the small effects of handedness on debitage characteristics with more power. Additionally, it would benefit researchers to explore alternate means of data analysis, instead of simply replicating current methods. Of particular importance is the application of more robust statistical techniques to existing data (see Ruck 2014a), and perhaps a shift toward more quantitative data collection methods as well.

The use of experimental archaeology to determine handedness in extinct hominids is a relatively new approach in paleoanthropology, but its implications are vast. Out of the approaches that form the basis of this study, only three (Toth 1985; Uomini 2001, 2006) applied experimental data to fossil evidence, with extremely mixed results. Still, the sheer breadth of lithic materials recovered from Lower and Middle Paleolithic deposits, especially in contrast to the dearth of other paleoarchaeological data, reflects a huge untapped resource in understanding the evolution of hominid handedness. The inherent weaknesses of this study and the methodologies reviewed are secondary to the benefits that future studies may provide regarding hominid cognitive and behavioral evolution. Paleoanthropologists interested in cognitive evolution need to direct their attention to the role of handedness in the fossil record, as it may be the best proxy for brain lateralization we have. Researchers interested in hominid handedness need to, in turn, focus more on lithic-based experimental approaches to inferring handedness, and acknowledge the best naturally preserved evidence our ancestors have left behind.

ACKNOWLEDGEMENTS

This project would not have been possible without the help of our volunteer flintknappers and raw materials suppliers. We would also like to thank Miki Matrullo and Katherine Sloate for their help in cataloging lithic materials. Special thanks to Justin Colón and Clifford Brown for assessing a random sample of flakes. Finally, we would like to thank Dr Natalie Uomini for her constant help with this research.

REFERENCES

Ambrose, S. H. 2001 Paleolithic Technology and Human Evolution. *Science* 291:1748–1753.

Annett, M. 1985 *Left, Right, Hand and Brain: The Right Shift Theory*. Lawrence Erlbaum Associates Ltd., London.

Annett, M. 1998 Stories about Hands, Brains, and Minds. *Brain and Language* 65:356–358.

Balzeau, A., E. Gilissen, and D. Grimaud-Hervé 2011 Shared Pattern of Endocranial Shape Asymmetries among Great Apes, Anatomically Modern Humans, and Fossil Hominins. *PLoS One* 7:e29581.

Bax, J. S., and P. S. Ungar 1999 Incisor Labial Surface Wear Striations in Modern Humans and Their Implications for

Handedness in Middle and Late Pleistocene Hominids. *International Journal of Osteoarchaeology* 198:189–198.

Bargalló, A., and M. Mosquera 2013 Can Hand Laterality Be Identified through Lithic Technology? *Laterality* 18:1–27.

Ben-David, A. 2008 Comparison of Classification Accuracy Using Cohen's Weighted Kappa. *Expert Systems with Applications* 34(2): 825–832.

Bermúdez de Castro, J., T. G. Bromage, and Y. F. Jalvo 1988 Buccal Striations on Fossil Human Anterior Teeth: Evidence of Handedness in the Middle and Early Upper Pleistocene. *Journal of Human Evolution* 17:403–412.

Bordes, F. 1968 *The Old Stone Age.* edited by J. E. Anderson. World University Library, New York.

Bromage, T. G., and A. Boyde 1984 Microscopic Criteria for the Determination of Directionality of Cutmarks on Bone. *American Journal of Physical Anthropology* 65:359–366.

Bruner, E. (editor) 2015 *Human Paleoneurology*. Springer, Switzerland.

Byrne, R. W. 2005 The Maker Not the Tool: the Cognitive Significance of Great Ape Manual Skills. In *Stone Knapping: The Necessary Conditions for a Uniquely Hominin Behaviour*, edited by V. Roux and B. Bril, pp. 159–170. McDonald Institute of Archeological Research, Cambridge.

Cantalupo, C., and W. D. Hopkins 2008 Theoretical Speculations on the Evolutionary Origins of Hemispheric Specialization. *Current Directions in Psychological Science* 17:233–237.

Cashmore, L. 2009 Can Hominin 'Handedness' Be Accurately Assessed? *Annals of Human Biology* 36:624–641.

Cashmore, L., N. Uomini, and A. Chapelain 2008 The Evolution of Handedness in Humans and Great Apes: A Review and Current Issues. *Journal of Anthropological Sciences* 86:7–35.

Coolidge, F. L., and T. Wynn 2009 *The Rise of Homo sapiens: The Evolution of Modern Thinking.* Wiley-Blackwell, West Sussex.

Corballis, M. C. 1983 *Human Laterality*. Academic Press, New York.

Corballis, M. C. 1991 *The Lopsided Ape: Evolution of the Generative Mind.* Oxford University Press, Oxford.

Corballis, M. C. 2003 From Mouth to Hand: Gesture, Speech, and the Evolution of Right-Handedness. *Behavioral and Brain Sciences* 26:199–260.

Cornford, J. M. 1986 Specialised Resharpening Techniques and Evidence of Handedness. In *La Cotte de St. Brelade: Excavations by C.B.M. McBurney*, edited by P. Callow and J. M. Cornford, pp. 337–362. Geo Books, Cambridge, UK.

Cotterell, B., and J. Kamminga 1990 Stone Tools. In *Mechanics of Preindustrial Technology*, edited by B. Cotterell and J. Kamminga, pp. 125–155. Cambridge University Press, New York.

Davidson, I. 2010 The Archeology of Cognitive Evolution. *Wiley Interdisciplinary Reviews Cognitive Science* 1(2): 214–229.

de Beaune, S. A., F. L. Coolidge, and T. Wynn 2009 *Cognitive Archeology and Human Evolution.* Cambridge University Press, New York.

Fagot, J., and J. Vauclair 1991 Manual Laterality in Nonhuman Primates: A Distinction Between Handedness and Manual Specialization. *Psychological Bulletin* 109:76–89.

Faurie, C., and M. Raymond 2004 Handedness Frequency over More Than Ten Thousand Years. *Proceedings of the Royal Society Series B: Biological Sciences* 271:S43–S45.

Fleiss, J. L., J. Cohen, and B. S. Everitt 1969 Large Sample Standard Errors of Kappa and Weighted Kappa. *Psychological Bulletin* 72(5): 323–327.

Fox, C. L., and D. W. Frayer 1997 Non-dietary Marks in the Anterior Dentition of the Krapina Neanderthals. *International Journal of Osteoarchaeology* 7:133–149.

Frayer, D. W., Lozano, M., Bermúdez de Castro, J., Carbonell, E., Arsuaga, J.-L., Radovcic, J., Fiore, I., and L. Bondioli 2012 More than 500,000 Years of Right-handedness in Europe. *Laterality* 17(1): 51–69.

Geribàs, N., M. Mosquera, and J. M. Vergès 2010 The Gesture Substratum of Stone Tool Making: An Experimental Approach. *Annali dell'Universita di Ferrara Museologia Scientifica e Naturalistica* 6:155–162.

Gibson, K. R., and T. Ingold (editors) 1993 *Tools, Language, and Cognition in Human Evolution.* Cambridge University Press, Cambridge.

Hewes, G. H. 1993 A History of Speculation on the Relation between Tools and Language. In *Tools, Language and Cognition in Human Evolution*, edited by K. R. Gibson and T. Ingold, pp. 20–32. Cambridge University Press, New York.

Holloway, R. L. 2008 The Human Brain Evolving: A Personal Retrospective. In *The Human Brain Evolving: Paleoneurological Studies in Honor of Ralph L. Holloway*, edited by D. Broadfield, M. Yuan, K. D. Schick, and N. Toth, pp. 1–14. Stone Age Institute Press, Gosport.

Hopkinson, T., and M. J. White 2005 The Acheulean and the Handaxe: Structure and Agency in the Paleolithic. In *The Hominid Individual in Context: Archaeological Investigations of Lower and Middle Paleolithic Landscapes, Locales, and Artefacts*, edited by C. Gamble and M. Porr, pp. 13–28. Routledge, New York.

Hopkins, W. D., K. A. Phillips, A. Bania, S. E. Calcutt, J. Russell, J. Schaeffer, E. Lonsdorf, S. R. Ross, and S. J. Schapiro 2011 Hand Preferences for Coordinated Bimanual Actions in 777 Great Apes: Implications for the Evolution of Handedness in Hominins. *Journal of Human Evolution* 60:605–611.

Hopkins, W. D., and J. K. Rilling 2000 A Comparative MRI Study of the Relationship Between Neuroanatomical Asymmetry and Interhemispheric Connectivity in Primates: Implication for the Evolution of Functional Asymmetries. *Behavioral Neuroscience* 114:739–748.

Hopkins, W. D., J. L. Russell, S. Lambeth, and S. J. Schapiro 2007 Handedness and Neuroanatomical Asymmetries in Captive Chimpanzees: A Summary of 15 Years of Research. In *The Evolution of Hemispheric Specialization in Primates*, edited by W. D. Hopkins, pp. 147–181. Elsevier, Norman.

Khreisheh, N., Davies, D., and B. Bradley 2013 Extending Experimental Control: The Use of Porcelain in Flaked Stone Experimentation. *Advances in Archaeological*

Practice: A Journal of the Society for American Archaeology 1(1): 37–46.

Lazenby, R. A. 2002 Skeletal Biology, Functional Asymmetry and the Origins of 'Handedness'. *Journal of Theoretical Biology* 218:129–138.

Liguria, G. 1997 A Case of Marked Bilateral Asymmetry in the Upper Limbs of an Upper Paleolithic Male from Barma. *International Journal of Osteoarchaeology* 7:18–38.

Lozano, M., M. Mosquera, J. M. Bermúdez de Castro, J. L. Arsuaga, and E. Carbonell 2009 Right Handedness of *Homo heidelbergensis* from Sima de los Huesos (Atapuerca, Spain) 500,000 Years Ago. *Evolution and Human Behavior* 30:369–376.

Matsuzawa, T. 2001 Primate Foundations of Human Intelligence: A View of Tool Use in Nonhuman Primates and Fossil Hominids. In *Primate Origins of Human Cognition and Behavior*, edited by T. Matsuzawa, pp. 3–25. Springer-Verlag, Hong Kong.

McGrew, W. C., and L. F. Marchant 1996 On Which Side of the Apes? Ethological Study of Laterality of Hand Use. In *Great Ape Societies*, edited by W. C. McGrew, L. F. Marchant, and T. Nishida, pp. 255–272. Cambridge University Press, New York.

McGrew, W. C., and L. F. Marchant 1997 On the Other Hand: Current Issues in and Meta-Analysis of the Behavioral Laterality of Hand Function in Nonhuman Primates. *Yearbook of Physical Anthropology* 40:201–232.

McPherron, S. P. 2000 Handaxes as a Measure of the Mental Capabilities of Early Hominids. *Journal of Archaeological Science* 27(8): 655–663.

Moore, M. W. 2011 The Design Space of Stone Flaking: Implications for Cognitive Evolution. *World Archaeology* 43:702–715.

Nichols, T. R., P. M. Wisner, G. Cripe, and L. Gulabchand 2011 Putting the Kappa Statistic to Use. *The Quality Assurance Journal* 13(3): 57–61.

Noble, W., and I. Davidson 1996 *Human Evolution, Language and Mind: A Psychological and Archaeological Inquiry*. Cambridge University Press, New York.

Nowell, A., and I. Davidson 2010 *Stone Tools and the Evolution of Human Cognition*. University Press of Colorado, Boulder.

Patterson, K., and J. P. Sollberger 1986 Comments on Toth's Right-Handedness Study. *Lithic Technology* 15:109–111.

Peresani, M., and R. Miolo 2012 Small Shifts in Handedness Bias during the Early Mesolithic? A Reconstruction Inferred from Microburin Technology in the Eastern Italian Alps. *Journal of Anthropological Archaeology* 31:93–103.

Pickering, T. R., and B. Hensley-Marschand 2008 Cutmarks and Hominid Handedness. *Journal of Archaeological Science* 35:310–315.

Pobiner, B. 1999 The Use of Stone Tools to Determine Handedness in Hominids. *Current Anthropology* 40:90–92.

Rilling, J. K. 2008 Neuroscientific Approaches and Applications within Anthropology. *American Journal of Physical Anthropology* 47:2–32.

Roux, V., and B. Bril 2005 *Stone Knapping: The Necessary Conditions for a Uniquely Hominin Behaviour*. McDonald Institute for Archaeological Research, Cambridge.

Ruck, L. 2014a Experimental Archaeology and Hominid Evolution: Establishing a Methodology for Determining Hominid Handedness in Lithic Materials as a Proxy for Cognitive Evolution. Unpublished M.A. thesis, Department of Anthropology, Florida Atlantic University, FL, USA.

Ruck, L. 2014b Manual Praxis and Stone Tool Manufacture: Implications for Language Evolution. *Brain and Language* 139C:68–83.

Rugg, G., and M. Mullane 2001 Inferring Handedness from Lithic Evidence. *Laterality* 6:247–259.

Schick, K. D., and J. D. Clark 2003 Biface Technological Development and Variability in the Acheulean Industrial Complex in the Middle Awash Region of the Afar Rift, Ethiopia. In *Multiple Approaches to the Study of Bifacial Technologies*, edited by M. Soressi and H. L. Dibble, pp. 1–30. The University of Pennsylvania, Philadelphia.

Schick, K. D., and N. Toth 2006 *Making Silent Stones Speak: Human Evolution and the Dawn of Technology*. Simon and Schuster, New York.

Semenov, S. A. 1964 *Prehistoric Technology: An Experimental Study of the Oldest Tools and Artefacts from Traces of Manufacture and Wear*, edited by M. W. Thompson. Cory, Adams, and Mackay Ltd, London, UK.

Shaw, C. N. 2011 Is 'Hand Preference' Coded in the Hominin Skeleton? An In-Vivo Study of Bilateral Morphological Variation. *Journal of Human Evolution* 61:480–7.

Sherwood, C. C., D. C. Broadfield, R. L. Holloway, P. J. Gannon, and P. R. Hof 2003 Variability of Broca's Area Homologue in African Great Apes: Implications for Language Evolution. *The Anatomical Record. Part A, Discoveries in Molecular, Cellular, and Evolutionary Biology* 271:276–85.

Steele, J. 2000 Handedness in Past Human Populations: Skeletal Markers. *Laterality* 5:193–220.

Steele, J., and N. T. Uomini 2005 Humans, Tools, and Handedness. In *Stone Knapping: The Necessary Conditions for a Uniquely Hominin Behaviour*, edited by V. Roux and B. Bril, pp. 217–255. McDonald Institute of Archeological Research, Cambridge.

Steele, J., and N. T. Uomini 2009 Can the Archaeology of Manual Specialization Tell Us Anything About Language Evolution? A Survey of the State of Play. *Cambridge Archaeological Journal* 19:97–110.

Stout, D. 2005 Neural Foundations of Perception and Action in Stone Knapping. In *Stone Knapping: The Necessary Conditions for a Uniquely Hominin Behaviour*, edited by V. Roux and B. Bril, pp. 273–286. McDonald Institute of Archeological Research, Cambridge.

Stout, D., and T. Chaminade 2012 Stone Tools, Language and the Brain in Human Evolution. *Philosophical Transactions of the Royal Society of London—Series B: Biological Sciences* 367:75–87.

Stout, D., and S. Semaw 2006 Knapping Skill of the Earliest Stone Toolmakers: Insights from the Study of Modern Human Novices. In *The Oldowan: Case Studies into the*

Earliest Stone Age, edited by N. Toth and K. D. Schick, pp. 307–320. Stone Age Institute Press, Gosport.

Stout, D., N. Toth, and K. D. Schick 2006 Comparing the Neural Foundations of Oldowan and Acheulean Toolmaking: A Pilot Study Using Positron Emission Tomography (PET). In *The Oldowan: Case Studies into the Earliest Stone Age*, edited by N. Toth and K. D. Schick, pp. 321–332. Stone Age Institute Press, Gosport.

Stout, D., N. Toth, K. D. Schick, and T. Chaminade 2008 Neural Correlates of Early Stone Age Toolmaking: Technology, Language and Cognition in Human Evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363:1939–1949.

Stout, D., N. Toth, K. D. Schick, J. Stout, and G. Hutchins 2000 Stone Tool-Making and Brain Activation: Position Emission Tomography (PET) Studies. *Journal of Archaeological Science* 27:1215–1223.

Toth, N. 1985 Archaeological Evidence for Preferential Right-handedness in the Lower and Middle Pleistocene, and Its Possible Implications. *Journal of Human Evolution* 14:607–614.

Toth, N., and K. D. Schick 1986 The First Million Years: The Archaeology of Protohuman Culture. *Advances in Archaeological Method and Theory* 9:1–96.

Trinkaus, E., S. E. Churchill, and C. B. Ruff 1994 Postcranial Robusticity in *Homo*. II: Humeral Bilateral Asymmetry and Bone Plasticity. *American Journal of Physical Anthropology* 93(1): 1–34.

Ubelaker, D. H., and L. M. Zarenko 2012 Can Handedness Be Determined from Skeletal Remains? A Chronological Review of the Literature. *Journal of Forensic Sciences* 57:1421–6.

Uomini, N. T. 2001 Lithic Indicators of Handedness: Assessment of Methodologies and the Evolution of Laterality in Hominids. Unpublished M.Sc. thesis, Department of Biological Anthropology, University of Durham, UK.

Uomini, N. T. 2006 In the Knapper's Hands: Testing Markers of Laterality in Hominin Lithic Production, with Reference to the Common Substrate of Language and Handedness. Unpublished Ph.D. dissertation, Department of Archaeology, University of Southampton, UK.

Uomini, N. T. 2009 Prehistoric Handedness and Prehistoric Language. In *Cognitive Archaeology and Human Evolution*, edited by S. A. de Beaune, F. L. Coolidge, and T. Wynn, pp. 37–56. Cambridge University Press, New York.

Uomini, N. T. 2011 Handedness in Neanderthals. In *Neanderthal Lifeways, Subsistence and Technology*, edited by N. J. Conard and J. Richter, pp. 139–158. Springer, Heidelberg.

Uomini, N. T. and G. F. Meyer 2013 Shared Brain Lateralization Patterns in Language and Acheulean Stone Tool Production: A Functional Transcranial Doppler Ultrasound Study. *PLoS One* 8(8): e72693.

Vandermeersch, B., and E. Trinkaus 1995 The Postcranial Remains of the Regourdou 1 Neandertal: The Shoulder and Arm Remains. *Journal of Human Evolution* 28:439–476.

Volpato, V., R. Macchiarelli, D. Guatelli-Steinberg, I. Fiore, L. Bondoli, and D. W. Frayer 2012 Hand to mouth in a Neandertal: Right-handedness in Regourdou 1. *PLoS One* 7:e43949.

Warrens, M. J. 2011 Cohen's Kappa Is a Weighted Average. *Statistical Methodology* 8(6): 473–484.

Wynn, T. 2002 Archaeology and Cognitive Evolution. *Behavioral and Brain Sciences* 25:389–438.

Notes on Contributors

Lana Ruck earned her M.A. in Anthropology at Florida Atlantic University, and will be a doctoral student in Cognitive Science and Anthropology and Indiana University—Bloomington starting in the Fall of 2015.

Douglas C. Broadfield is an Associate Professor in the Department of Cell Biology at the University of Miami—Miller School of Medicine. His current research includes human plaeoneurology, primate comparative neurology, and language evolution. His most recent book is *The Human Brain Evolving*.

Clifford T. Brown is an Associate Professor in the Department of Anthropology at Florida Atlantic University. He studied archaeology at Yale and anthropology at Tulane. In between bouts of field work in Mexico and Nicaragua, he conducts research on quantitative methods in archaeology.

Correspondence to: Lana Ruck, 500 N. Congress Ave., Apt. D-102, Delray Beach, FL 33445, USA. Email: lruck@fau.edu