Lithic Technology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ylit20

Skill and Core Uniformity: An Experiment with Oldowan-like Flaking Systems

Evan Patrick Wilson, Dietrich Stout, Cheng Liu, Megan Beney Kilgore & Justin Pargeter

To cite this article: Evan Patrick Wilson, Dietrich Stout, Cheng Liu, Megan Beney Kilgore & Justin Pargeter (2023) Skill and Core Uniformity: An Experiment with Oldowan-like Flaking Systems, Lithic Technology, 48:4, 333-346, DOI: 10.1080/01977261.2023.2178767

To link to this article: https://doi.org/10.1080/01977261.2023.2178767

LITHIC TECHNOLOGY
2023, VOL. 48, NO. 4, 333–346
https://doi.org/10.1080/01977261.2023.2178767

Skill and Core Uniformity: An Experiment with Oldowan-like Flaking Systems

Evan Patrick Wilson[a], Dietrich Stout[b], Cheng Liu[b], Megan Beney Kilgore[b] and Justin Pargeter[c,d]

[a]Department of Anthropology, City University of New York Graduate Center, New York, NY, USA; [b]Department of Anthropology, Emory University, Atlanta, GA, USA; [c]Department of Anthropology, New York University, New York, NY, USA; [d]Rock Art Research Institute, School of Geography, Archaeology, and Environmental Sciences, University of the Witwatersrand, Johannesburg, South Africa

ABSTRACT

The Oldowan is the archaeological record's oldest consistent evidence of hominin technical behavior. First appearing ∼2.6 Ma in East Africa, the Oldowan is characterized by simple core and flake technology using direct hard hammer percussion. Archaeologists debate whether Oldowan assemblages are uniform and what role hominin cultural abilities played in generating these assemblages. To improve existing methods for studying Oldowan technical uniformity, we conducted experiments involving 23 novices and one expert knapper. Subjects made simple stone tools under two different instructional conditions (observation-only and direct active instruction) over two hours. We used the resulting cores to track flaking efficiency, reduction intensity, and knapping errors. We find significant differences in the expert and novice core uniformity. Direct active teaching increased core flaking efficiency and reduced knapping errors. Comparisons between our experimental results and an Oldowan sample from Gona, Ethiopia, show core variability patterns that match our expert and actively taught novices.

ARTICLE HISTORY
Received 20 June 2022
Accepted 6 February 2023

KEYWORDS
Oldowan; stone tool-making; social learning; uniformity

Introduction

Archaeologists have done relatively little experimentation to understand the relationship between skill and standardization in the production of stone tool technology. Although often discussed, both standardization and skill tend to be variably defined (Bamforth & Finlay, 2008; Eerkens & Bettinger, 2001; Kuhn, 2010; Monnier & McNulty, 2010), which can create confusion in the literature given that archaeologists regularly equate technical skill with technological standardization (Bamforth & Finlay, 2008; Lassen & Williams, 2015). This ascribed equivalency is understandable, given that artefacts may be more standardized when knapped via a series of highly repetitive action sequences reliant on the combined and increased mastery of abstract knowledge and motor-skill know-how (Pelegrin, 1990). However, it is important to note the differences in definitions of skill and standardization and how these differences may affect the interpretation of experimental results.

In this study we define standardization as a reduction in variability (Marks et al., 2001). For stone products, standardization typically refers to reduced variability in artefact shapes and sizes. A less frequent, but no less important, definition of standardization refers to reduced variability in how toolmakers make stone tools (Kuhn, 2010). Stone tool standardization manifests differently in core and flake assemblages such that standardized procedures visible on cores might result in highly variable end products (flakes), such as is the case with bipolar flaking (Pargeter & de la Peña, 2017), and vice versa in the case of intensely shaped, standardized flaked tool forms. Because it is most often discussed in the context of highly standardized artefact types (e.g. Late Acheulean handaxes or Clovis Points) there is often an inherent link assumed between standardization, mental templates, and intention. However, it is equally as likely that honed sensorimotor control and know-how acquired through teaching and practice can impact the uniformity of stone cores and flakes in the absence of intentional efforts to standardize these artefacts (see Harush et al., 2020 and Harush & Grosman, 2021 for an example from ceramic production). For this reason, we focus the current paper on tracking the role of skill and training on stone core uniformity. Tracking skill acquisition's role in stone tool uniformity is especially important for the earliest parts of the archaeological record. Particularly so prior to 2 million years ago where uniform tool forms appear first and archaeologists have argued skillful stoneworking behaviors may have been present (Delagnes & Roche, 2005; Roche et al., 1999; Semaw et al., 1997).

In this work we explore uniformity in Oldowan style "simple" stone cores knapped by novice and expert knappers. We ask two primary questions:

CONTACT Justin Pargeter justin.pargeter@nyu.edu 25 Waverly Place, New York, NY, 10012, USA

1. Do experts and novices display significantly different degrees of core technological uniformity?
2. Do novices with direct-active teaching show increased core uniformity compared to novices provided with observation-only training?

We recorded and compared aspects of reduction intensity, flake invasiveness, core battering, and core geometry following an attribute system to compare technical skill in Oldowan assemblages. Our study compared the features of cores produced by 23 novices and one expert during a two-hour training period. We find that significant differences in skill are apparent between our expert and novice knappers and that the degree of variability and uniformity in their end-products is also significantly different. This is suggestive of skill-driven variance. We find significant effects of direct active teaching on increased core flaking invasiveness and reduced core edge battering. Our findings are in line with Morgan et al. (2015) conclusions from similar experimental work, which highlighted the importance of active teaching and language in the transmission of stone knapping skills. Comparisons between our experimental results and an Oldowan sample from Gona, Ethiopia, shows intermediate archaeological core variability that matches well with our expert and actively taught novices.

Skill and lithic technology

Archaeologists argue that varying skill-level in the production of stone tools is an important factor in the creation of lithic technological variation (Bamforth & Finlay, 2008; Pargeter et al., 2019). Yet, the acquisition and archaeological manifestation of knapping skill have received relatively little structured scholarly attention and thus remain difficult to characterize (Bamforth & Finlay, 2008; Herzlinger et al., 2017; Pargeter et al., 2019, 2020, 2022; Torres & Preysler, 2020). As Bamforth and Finlay (2008) note, skill is a concept that defies any single definition, spans a wide variety of situations, and is highly dependent on context. As such, teasing out information relevant to the relationship between skill and uniformity from the material record is a particularly multiplex problem. Despite these incongruities, scholars appear to be in general agreement that skill involves the practical application of abstract knowledge in a goal-directed manner which can be improved with practice and training (Bamforth & Finlay, 2008; Herzlinger et al., 2017; Ingold, 1993; Olausson, 2008; Proctor & Dutta, 1995), as such this will serve as our definition of "skill" for the purposes of this paper.

Initial studies of knapping skill were largely descriptive and based on subjective criteria. These studies often focus on replicating technological systems and end products with a high degree of morphological standardization, such as Danish flint daggers (Apel, 2001) or fluted Paleoindian points (Flenniken, 1978). Such standardized technologies are also often used in learning and cultural transmission studies (Apel, 2008; O'Brien et al., 2014). However, stone tool making is a complex, goal-oriented behavior in which variability is manifest at several points in the production phase and not just in the final tools. Elementary success requires simultaneous control over various interacting variables, including the force of a blow, the angle of an impact, and the stabilization of the core during knapping (Bleed, 2008; Bril et al., 2010; Roux & David, 2005). Achieving higher levels of skill requires even more consistent control over these variables gained through continued deliberate practice and training (Bleed, 2008; Ericsson et al., 1993; Olausson, 2008; Pargeter et al., 2019). Mastery of stone knapping observed in modern experts and ethnographic contexts can take years and is highly dependent on continued experience, often within structured social settings (Bleed, 2008; Ericsson et al., 1993; Stout, 2005; Stout et al., 2002), or "communities of practice" (Lave & Wenger, 1991; Wendrich, 2013). This continued practice and eventual mastery is associated with a knapper's ability to achieve higher degrees of both procedural and morphological uniformity even when starting with highly variable raw materials.

Archaeologists typically distinguish between knowledgeable practice (also called *connaissance*) and practical knowledge (also called *savoir-faire*) (sensu Pelegrin, 1993), the two broad realms of knowledge employed during knapping. Knowledgeable practice refers to the abstract planning process involved in stone tool making, including "cognitive understanding" of the techno-behavior being undertaken and "strategic decision making" (Bamforth & Finlay, 2008; Pargeter et al., 2020). Practical knowledge encompasses the practical aspects of knapping, or know-how, including manual dexterity and technical actions (Bamforth & Finlay, 2008; Pelegrin, 1990). Skill can be considered the active engagement and interaction between these two general realms of knowledge at their intersection (Bamforth & Finlay, 2008. Bril et al., 2010). A mastery of both realms and the ability to employ them simultaneously is essential for increasing flaking predictability and the consistent production of more standardized end-products (Nonaka et al., 2010).

Through continued practice the relationship between these knowledge domains will change and develop, as will the relation between the toolmaker's mind and the material being worked (Bamforth & Finlay, 2008; Olausson, 2008). Although sociocultural context is a major.

factor affecting skill, individual skill is largely determined by some (poorly understood) combination of the amount of time devoted to practicing/training and individualities such as baseline aptitude (Eren et al., 2008; Herzlinger et al., 2017; Olausson, 2008).

Skill and uniformity in the early stone age

Knapping skill and uniformity in the earliest stages of stone tool evolution have been of long interest to archaeologists and paleoanthropologists (Guilmet, 1977; Isaac, 1986; Oakley, 1969; Roche et al., 1999; Stout et al., 2009; Toth et al., 2006). This is, at least in part, due to preconceived notions long held in Western science and society about linear technological progress throughout human history and its link to our cognitive and cultural evolution (Monnier & McNulty, 2010; Oakley, 1969; Trigger, 1996). Despite this interest, a much larger proportion of work focused on uniformity in lithic technology has been devoted to Middle/Upper Paleolithic or Middle/Later Stone Age transitions than earlier periods (Marks et al., 2001; Monnier & McNulty, 2010), perhaps due to the perceived technological stagnation and lack of uniformity in early stone technologies, such as the Oldowan.

The Oldowan Industrial Complex (OIC) is the oldest well-known and best-documented evidence for hominin behavior found in the archaeological record (but see Harmand et al., 2015) and represents a major dietary, adaptive shift for our lineage (Barsky, 2009; Plummer, 2004; Toth & Schick, 2018). First appearing ~2.6 Ma in East Africa (Braun et al., 2019; Semaw et al., 1997, 2003), the Oldowan is characterized as a relatively simple core and flake technology directed towards the production of sharp flakes via hard hammer percussion (Barsky, 2009; de la Torre, 2004; Leakey, 1971; Plummer, 2004; Toth & Schick, 2018). Assemblages classified OIC appear until about 1 Ma (Toth & Schick, 2018) but Oldowan-like, or "Mode 1" (Clark, 1969), lithic technology appears in Eurasia even more recently in time (Barsky 2009).

The end-products of Oldowan knapping are considered largely non-uniform throughout time and lack any imposed tool forms (Braun et al., 2019; de la Torre, 2004; Plummer, 2004; Sahnouni et al., 2002; Semaw et al., 1997). Some recent work from younger OIC sites like Garba IV at Melke Kunture, Ethiopia, dated to ~1.7 Ma, has shown potential evidence for imposed form and artefact uniformity, in this case associated with the exploitation of high quality obsidian for the production of small pointed objects (Gallotti & Mussi, 2015). The authors argue that this does not require the invocation of a "cognitive leap," but supports the capacity for a degree of uniformity in the OIC.

While evidence for morphological uniformity and imposition of shape is scant in the Oldowan, some scholars have identified relatively sophisticated knapping capabilities and tool behaviors at some sites, even some of the earliest known (Delagnes & Roche, 2005; Finestone, 2019; Roche et al., 1999; Semaw et al., 1997). At sites such as the 2.3 million year old Lokalalei 2C in West Turkana, lithic refits, in some cases nearly entire reduction sequences, demonstrate early Oldowan knappers were able to control knapping in an efficient, regular, and structured manner (Roche et al., 1999). Later Oldowan sites have been argued to bear evidence for increased technical abilities and in some cases "technological homogeneity" (Sánchez-Yustos et al., 2017), perhaps anticipating subsequent Acheulean technology. These sites are sometimes classified as Developed Oldowan (Braun et al., 2008; de la Torre, 2011; de la Torre & Mora, 2005, 2014, Proffitt, 2018; Sánchez-Yustos et al., 2017; Stiles, 1979), but more recently the consensus has shifted away from this classification.

These inconsistent patterns of skill and uniformity across both space and time have been a matter of great discussion in the literature. Some hold out that, despite the observed range of variation, there is a general trend to increased skill and technological complexity in the OIC through time (Carbonell et al., 2016; de Lumley, 2022; de Lumley et al., 2009; Sánchez-Yustos, 2021). This is despite mounting evidence for relatively high levels of skill and complex knapping behavior early in the Oldowan (Delagnes & Roche, 2005; Finestone, 2019; Roche et al., 1999; Semaw et al., 1997), comparably lackluster exhibitions of skill at some more recent sites like HWK EE (~1.7 Ma) (de la Torre & Mora, 2018) and Kilombe (~1.78 Ma) (Gowlett et al., 2022), and nearly opposite levels of skill displayed at nearly the same time within the same localities (e.g. Lokalalei 1A and 2C; Delagnes & Roche, 2005; Kibunjia 1994; Roche et al., 1999).

Toth et al. (2006) attempted to evaluate skill in the Oldowan using an extensive list of attributes for both flakes and cores with the potential to indicate knapper skill. They compared Oldowan-style artefacts experimentally produced by modern-day *Homo sapiens* and trained captive *Pan paniscus*, the bonobo Kanzi and his half-sister Panbanisha, as well as archaeological material produced by hominins at the 2.6 million year old Oldowan site of Gona, Ethiopia (Semaw et al., 1997). Humans were used as the expert group, Oldowan hominins as the intermediate, and bonobos as the lower skill level. This study found that many of the authors' recorded attributes were not significantly different between their different skill groups, narrowing down their list to a series of “key” attributes. These are related to overall core reduction efficiency (i.e. more cutting edge per unit mass of stone and more invasive flaking) and maintenance of the platform edges (i.e. less edge-battering, suitable platform angles).

These authors found that *Homo sapiens* were better at reducing cores and limiting mistakes/knapping accidents. Gona hominins were found to have relatively sophisticated abilities for core reduction and the efficient production of flakes, perhaps not as efficient as the modern knappers, but more so than the bonobo assemblage. This assemblage was either clustered with *H. sapiens* or was intermediate between humans and bonobos, depending on the particular attribute. While *P. paniscus* was able to produce stone artefacts through conchoidal fracture not totally unlike the Oldowan, their assemblages bore a number of distinctions that clearly separated them from Oldowan toolmakers. These include the production of less flakes and relatively limited core reduction as well a greater degree of variability in their ability to avoid edge battering and maintain acute platform angles. Recent work suggests that Kanzi’s poor training environment and lack of sufficiently scaffolded learning may have led to his reduced flaking performance (Eren et al., 2020). These results support the need for further investigation into the role of teaching, training, and skill acquisition on Oldowan-like stone toolmaking capacities.

Materials and methods

In addition to one expert knapper acting as a demonstrator and instructor, a total of 23 healthy individuals lacking prior stone tool making experience were recruited from Emory University and the surrounding area for a pilot study on teaching Oldowan-style, least effort stone toolmaking. The goal of this experiment was to evaluate the effect of active expert instruction on novice knappers’ ability to produce flakes. Participants were randomly placed under one of two experimental conditions: observation-only or direct-active teaching. In the former condition the novice knappers were permitted to observe the demonstrator as well as other novices around them but could not discuss the task at hand and did not receive any direct instruction from the expert. Individuals in the latter group were unconstrained in the degree of communication permitted between participants and received direct instructions, including verbal and gestural, from the expert demonstrator (the second author).

Table 1. Relevant blank and core metrics for novice participants. Only metrics relevant to this study are reported and only for the novices.

	Mean	Median	SD	Min	Max
Starting mass					
Observation only ($n = 99$)	995.7	949.6	278.3	503.3	1876.2
Actively taught ($n = 108$)	956.2	862.5	296.9	459	1786.7
Starting max length					
Observation only ($n = 99$)	144.2	143.5	17.1	105.3	180.5
Actively taught ($n = 108$)	143.6	143	16.2	111	180.2
Final mass					
Observation only ($n = 99$)	236.8	180.5	179.4	28.5	953.9
Actively taught ($n = 108$)	144.8	128.6	89.9	17.5	371
Final max length					
Observation only ($n = 99$)	82.1	76.1	21.7	43	169.2
Actively taught ($n = 108$)	70.3	68.7	16.2	36.6	110.9

Knapping protocol

Each individual was supplied 9 nodules (see Table 1 for start and end metrics for blanks/cores) of the same variety of basalt and allowed to produce flakes via freehand direct percussion over the course of a 2-hour period. While the majority of flaking attempts were freehand, some participants, mostly in the observation-only group, spontaneously adopted a bipolar technique using the ground as an anvil (see Pargeter et al., In Press for further details). Raw material was kept consistent across knappers to control for the effects of raw material variability, a factor which has been found by Proffitt et al. (2022) to greatly affect both the manifestation of varying skill levels in the material record as well as inhibit the identification of skill variation. Furthermore, this kind of raw material is archaeologically consistent with materials used by early tool making hominins for Oldowan stone tool manufacture (Plummer, 2004). Participants chose hammerstones used for flaking from a provided selection which allowed them to choose stones that were both ergonomic and comfortable to use. Under both instructional conditions, novices had a chance to examine their intended end-product, flakes, prior to knapping and were informed that the desired end-goal was to produce as many flakes as possible within the given time. Cores were considered exhausted when participants felt they could no longer reduce them. A single expert knapper was used to reduce inter-instructor bias.

Target analysis

We analyzed all cores produced by both the expert and novices in both instruction conditions throughout the two-hour knapping session ($n = 190$). This allowed us to not only compare the effects of skill, but also evaluate the effect of training condition for each novice. Only cores were selected for this analysis as we report on

Table 2. All recorded attributes, their definitions, how they were measured, and our predictions for each attribute with increasing uniformity.

Attribute	Definition	How we measured the variable	Predictions under Increased uniformity
Flake scar density index (SDI)	A measure of reduction intensity capturing a knappers ability to remove more flakes per core unit mass. Higher values show greater reduction intensity.	All flake scars >10 mm were counted and divided by the surface area of the core generated from 3D scans.	With increased uniformity we predict the average SDI to be similar across an individual's cores. The expert is expected to produce higher SDI values than novices.
Percentage mass reduced	Final core mass as a percentage of original nodule mass. Higher values show greater overall reduction intensity.	Cores were weighed before and after knapping on a scale.	Experts should show greater amounts of mass reduced on their cores more consistently. Novices should be more variable in their overall reduced mass.
Percentage Remaining Cortex	A measure of a knappers ability to remove the core's original outer surface and access the higher quality material within	Visually assessed to the nearest 5%	Increased uniformity should restrict the variability in remaining cortex frequency across cores indicating the removal of a consistent proportion of the core's original surface.
Flake Invasiveness	A measure of a knapper's ability to produce relatively long flakes using as much of the available flaking surface as possible. Higher values show greater differences between core maximum dimension and flake maximum dimension.	The maximum dimension of all flake scars >10 mm were averaged for each core and regressed against the maximum dimension. Residuals from this model were compared between the groups to eliminate allometric effects of core size differences.	Experts should be able to make flakes closer to the maximum dimension of the core more consistently than novices, leading to lower residual values.
Platform Battering	The frequency and extent of crushing and battering across the area of available striking platform. Captures a knappers failed removals, poor platform preparation, and misguided perseverance.	Visually assessed to the nearest 5%	With increased uniformity, the variability in knapping damage should decrease as knappers better grasp their own limits as well as the constraints of their raw material, retiring a core before battering it profusely.

the experiment's flake assemblage in a separate study (Pargeter et al., In Press). Cores were analyzed and evaluated for a number of attributes others have argued are key to evaluating flaking success in relatively simple Oldowan-like flaking systems (e.g. Stout et al., 2010, 2019; Toth et al., 2006). Overall, these attributes relate to an expert's ability to consistently reduce cores to a greater degree than novices by removing a higher number of relatively more invasive flakes and retain a flakable platform by minimizing edge battering while limiting knapping mistakes (Table 2).

We evaluated reduction intensity in three independent ways: (1) by the flake scar density index (Clarkson, 2013), (2) by comparing nodule starting mass to core final mass as a percentage of mass lost during the reduction, and (3) through a visual estimate of the percentage cortex remaining on the surface to the nearest 5%. Visual assessment has been found to be a reliable and effective method for discerning the percentage of remaining cortex on a core by Dibble et al. (2005). We tracked failed removals, poor platform preparation, and misguided perseverance by tracking the frequency and extent of battering across the area of available platform. A lower prevalence of platform battering was associated with more expertly flaked cores, so the percentage area of the available striking platform edge exhibiting battering was estimated through visual assessment also to the nearest 5%.

Statistical methods

Participants were placed in one of three experimental groups: expert, actively taught novices, and observation-only novices. Each resulting core variable in Table 2 was first plotted individually and compared via ANOVA to evaluate significant differences in the average performance between the three groups. For the flake scar invasiveness index, we regressed core maximum flake dimension against core maximum dimension to derive a set of residuals that accounted for this relationship after accounting for allometric effects. We then compared these residuals between the three groups. Cortex values are expressed as proportions (range: 0–1) and we modeled them using a beta regression using R version 4.0.3's (R Core Team, 2013) *betareg* package (Cribari-Neto & Zeileis, 2010). To determine if these skill-related attributes also track uniformity, we then compared our groups using a Bartlett test of homogeneity of variance to assess significant differences in variance in our recorded attributes. The paper's R code and raw datafiles are freely available here: https://github.com/Raylc/PaST_core_standardization.

Results

Group level comparisons between expert and novice knappers

The expert, on average, produced significantly higher flake scar densities ($F\ [2, 186] = 19.1$, $p < 0.01$) than

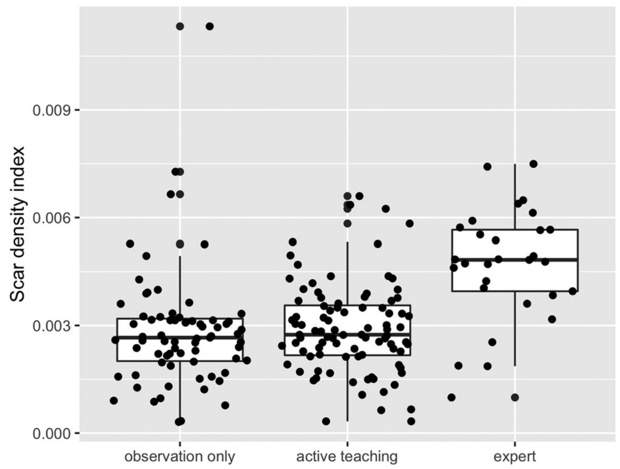

Figure 1. Comparison of the flake scar density index across our three groups. The expert shows significantly higher SDI values ($p < 0.01$) than novices with non-significant differences in variance ($p = 0.05$).

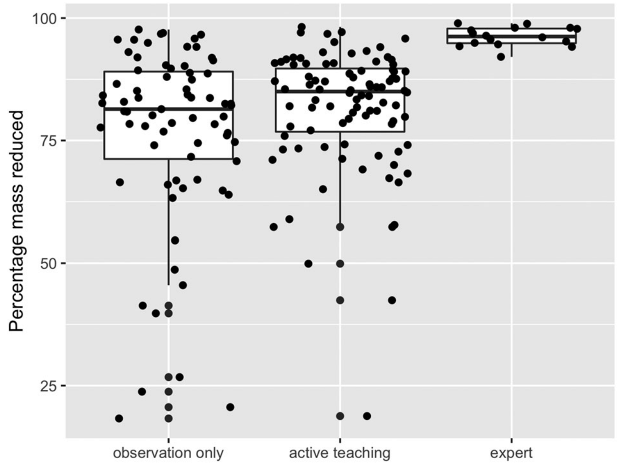

Figure 2. Comparison of percentage core mass reduced across our three groups. The expert reduced significantly more mass on their cores ($p < 0.01$) than both groups of novices with a significantly lower overall variance ($p < 0.01$).

novices in either teaching condition (Figure 1, Table 3). While on average higher, the Bartlett test indicates the expert exhibits the same range of variation (K-squared = 6.1, df = 2, p-value = 0.05) in the flake scar densities across their cores. This is unexpected and runs counter to our predictions of reduced variation in the expert sample.

Comparing the percentage of core mass reduced we find significant (F [2, 176] = 11.4, $p < 0.01$) differences between the expert and novice groups (Figure 2, Table 4). The two novice groups do not show significant differences to each other. Bartlett's test found significant differences in variance (K-squared = 58.2, df = 2, p-value < 0.01) between the groups driven by the expert's very low variance. The expert exhibits a more constrained range of variation than novices indicating their ability to regularly reduce cores to a greater degree. This result suggests that the SDI measure is not a precise measure of reduction intensity, which might justify the surprising result that our expert knapper showed similar SDI variance to our novices.

In the case of flake invasiveness we also find significant (F [2, 190] = 34.1, $p < 0.01$) average differences across our groups (Figure 3, Table 5). The expert once again plots away from both novice groups, producing relatively longer, more invasive flakes on average, reflecting their ability to regularly exploit more of the available flaking surface. The two training groups show significant differences Bartlett's test found significant differences in variance (K-squared = 46.9, df = 2, p-value < 0.01) across our three groups. The expert exhibits a more constrained range of variation than novices indicating their ability to regularly produce more invasive flakes.

We found significant differences in core edge battering frequencies between the groups (Figure 4, Table 6). The expert produced significantly less battering damage on average (F [2, 187] = 24.5, $p < 0.01$) than novices in both teaching conditions. The observation-only novices showed significantly higher average frequencies of edge battering compared to both experts and actively taught novices. Bartlett's test found significant differences in variance (K-squared = 65.4, df = 2, p-value < 0.01) across groups. The expert shows far less variation in edge battering compared to novices, perhaps reflecting his ability to more consistently apply the force necessary and in the proper place on the core to produce flakes. Actively taught novices sit intermediate between the expert and observation-only novices.

Figure 5 (also see Table 7) compares the cortex proportions across our three groups. The beta regression results show strong significant differences between the expert and novice groups (Phi = 4.1, $p < 0.01$). The two novice groups show no significant differences. The

Table 3. Scar Density Index (SDI) summary statistics for our three experimental groups.

	Mean	Median	SD	Min	Max
Observation only ($n = 71$)	.003	.003	.002	.000	.01
Actively taught ($n = 92$)	.003	.003	.001	.000	.007
Expert ($n = 29$)	.005	.005	.002	.001	.007

Table 4. Percentage of Core Mass Reduced summary statistics for our three experimental groups.

	Mean	Median	SD	Min	Max
Observation only ($n = 71$)	76.6	81.4	18.7	18.3	97.7
Actively taught ($n = 92$)	81.5	85	12.7	18.8	98.2
Expert ($n = 16$)	96.2	96.2	.484	92.1	98.9

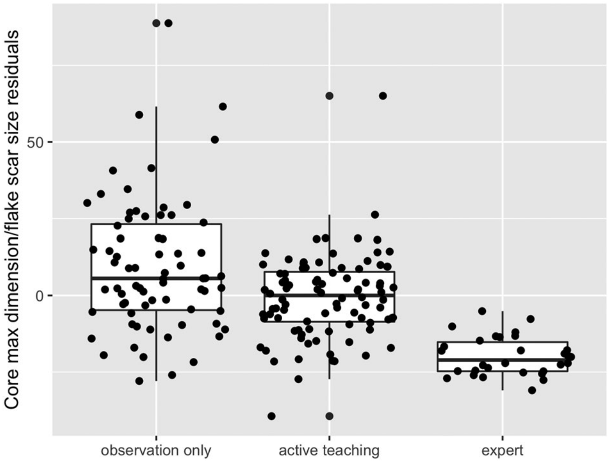

Figure 3. Comparison of flaking invasiveness across our three groups. The expert produced significantly more ($p < 0.01$) invasive flakes than novices with a significantly lower degree of variance ($p < 0.01$).

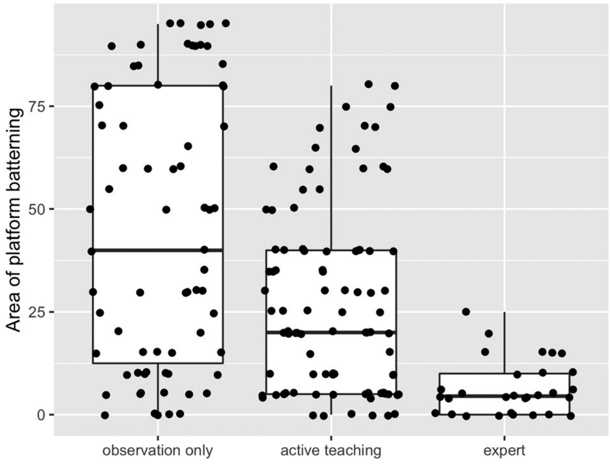

Figure 4. Comparison of platform battering, measured as the percentage of available platform area with battering, across our three groups. The expert battered their cores significantly less ($p < 0.01$) than both groups of novices with a significantly smaller degree of variance ($p < 0.01$).

expert showed a notably lower cortex standard deviation (sd = 12) compared to the observation only and actively taught novice groups (sd = 21 and sd = 18). This result demonstrates that the expert overall managed to reduce the nodule cortex more intensively and uniformly than the novices.

Comparisons with archaeological Oldowan cores from Gona, Ethiopia

Our archaeological comparisons are limited by the availability of comparable metrics from Oldowan sites. Toth et al. (2006) report summary data for core edge battering and surface cortex on cores from two sites with nearly identical lithic technology from the early Oldowan locality of Gona, Ethiopia, East Gona (EG) 10 and EG 12. The earliest deposits at Gona date to roughly 2.6 million years old, making it one of the oldest currently known Oldowan sites (Semaw et al., 2003). The combined EG10 and EG12 assemblage is dominated by unifacial flaked cores characterized by evidence for skillful core reduction efficiency and flake production (Toth et al., 2006). These data allow for limited qualitative comparisons with our experimental results.

Toth et al. (2006) scored edge battering in four classes: none, low, moderate, and high. The frequencies of battering in each of these classes showed a wider range in their Gona cores compared with their two expert knappers (Table 8). We recorded edge battering as a continuous variable, but one could class our data into four sub-portions based on edge battering frequencies (0%, <25%, 25–75%, 76–100%). The Gona data sit between our direct-actively taught and expert groups in having few to no instances of high edge battering, but with variance across the remaining three edge battering classes. Our observation-only novice group showed the inverse pattern with no cores absent edge battering and variance spread across the remaining three higher battering classes.

Toth et al. (2006) scored flake invasiveness as shallow, moderate and invasive. Their Gona and expert knapper samples show similar variances across these categories though with inverse patterns (higher invasive flaking in the Gona sample compared to their experts due to the nature of their experimental design) (Table 8). We reclassed our flaking invasiveness measure into three equally sized groups for comparisons with Toth et al. (2006) shallow, moderate and invasive classes. This result lines up with the Toth et al. (2006) data showing that Gona flaking invasiveness classes resemble our expert and actively taught groups more than the direct observation novice group.

Table 5. Flake Invasiveness summary statistics for our three experimental groups.

	Mean	Median	SD	Min	Max
Observation only ($n = 71$)	9.35	5.56	21.4	−27.9	88.8
Actively taught ($n = 92$)	−.743	−.005	13.7	−39.4	65
Expert ($n = 30$)	−19.8	−21.1	6.39	−31	−5.13

Table 6. Edge battering summary statistics for our three experimental groups.

	Mean	Median	SD	Min	Max
Observation only ($n = 73$)	38.8	35	29.7	0	95
Actively taught ($n = 87$)	32.8	25	28	0	95
Expert ($n = 30$)	6.4	4.5	6.6	0	25

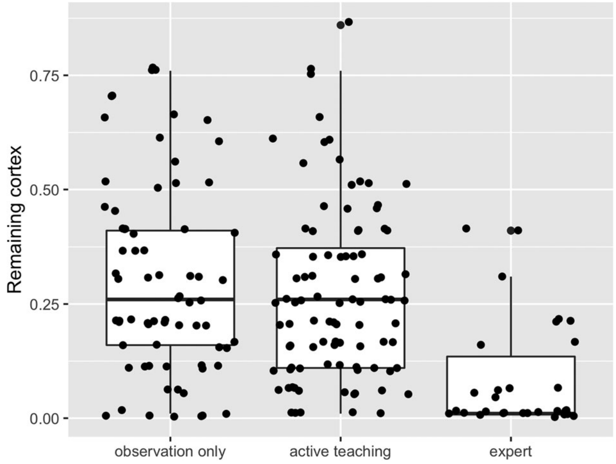

Figure 5. Comparison of core cortex across our three groups. The expert reduced significantly more cortex ($p < 0.01$) than both groups of novices with a notably more uniform distribution.

Toth et al. (2006) note higher variance in their Gona core cortex frequencies compared to their expert knappers (Mean/SD Gona = 53.4/16.1 vs. Mean/SD Expert = 61.6/11.2) (also see Stout et al., 2010: Supplementary Table 1). Our experiment showed a similar lower degree of expert core cortex variance (Mean/SD Expert = 7.43/12) compared to our novice samples (Mean/SD Novice = 27.5/19.7) with variance decreasing with direct active teaching (Mean/SD Direct active teaching = 25.9/18.3, Mean/SD observation-only = 29.6/21.2). This result suggests that Gona knappers managed core reduction intensity and cortex removal in a manner comparable to our expert and direct actively taught knappers and less like our observation-only group.

Although limited in number and scope due to incomparable data, the comparisons we can make between our results, Toth and colleagues' experiment, and the archaeological data indicate that Oldowan hominins achieved variable levels of skill across space and through time. Earlier Oldowan toolmakers at Gona produced flakes of similar invasiveness to both Toth et al.'s (2006) and our expert and managed reduction intensity and cortex removal more like our expert and actively taught novices than the observation-only group.

Table 7. Cortex frequency summary statistics for our three experimental groups.

	Mean	Median	SD	Min	Max
Observation only ($n = 73$)	33.5	30	19.2	0	85
Actively taught ($n = 87$)	22.5	20	18.7	0	75
Expert ($n = 30$)	7.4	0	12	0	40

Discussion

Our results show significant group-level differences in flake scar density, percentage mass lost during reduction, flake invasiveness, platform battering, and core cortex between our expert and the novice groups. These data reflect these groups' differing abilities to invasively flake and reduce cores efficiently. Our results also point to important effects of direct-active teaching with the actively taught novices showing significantly more invasive flaking and lower core edge battering than the observation only group, at least in this relatively short practice period. Given more practice time, it is possible these differences would disappear with more time for individual learning. A series of Bartlett tests of homogeneity of variances and examinations of cortex standard deviation found strong differences in four out of our five variables (scar density index did not show significant variance differences). As we predicted, the expert reduced more mass, flaked more invasively, showed less battering, and removed cortex more uniformly than either of the novice groups. Actively taught individuals were more uniform in their flaking invasiveness and edge battering than subjects in the observation-only group. This suggests that skill is significantly correlated with core uniformity in relatively simple Oldowan-style flaking and that the attributes recorded in this study can be used to track these differences.

Our expert shows a non-significant difference in the flake scar density index compared to both novice groups. This is counter to our initial predictions of reduced variance in reduction intensity due to more consistent reduction strategies and skillful applications of force. This is not completely surprising as during flaking, previous flake scars are often removed, thus, extensively exhausted cores will lack many, if not most, of the total flake scars produced. This interpretation is corroborated by the degree of mass reduced results which show clear and significant expert-novice differences.

Flaking invasiveness matched our predictions with the expert showing a significantly lower range of variability in the average invasiveness of their flakes than novices. The expert regularly exploited more of the available flaking surface than novices, hypothetically optimizing the amount of cutting edge produced from the available raw material. Unlike our reduction intensity measures, the actively taught novice groups made more invasive flakes compared with the observation-only group. Flaking invasiveness is particularly relevant to the archaeological record as, based on our data, it reliably tracks skill-related variation in core morphology

Table 8. Comparable attribute data from Toth et al. (2006) compared to data from this study.

		Gona		Toth experts		This study expert		Active teaching		Observation only	
		n	%	*n*	%	*n*	%	*n*	%	*n*	%
Edge battering	None	7	30.4	26	83.9	9	30	5	5.5	5	7.1
	Low	10	43.5	4	12.9	20	66.6	40	44.4	20	28.5
	Moderate	6	26.1	1	3.2	1	3.3	43	47.7	26	37.1
	High	0	0	0	0	0	0	2	2.2	19	27.1
Flaking invasiveness	Shallow	0	0	1	3.2	0	0	27	29.3	37	52.1
	Moderate	7	30.4	19	61.3	2	6.6	43	46.7	19	26.7
	Invasive	16	69.6	11	35.5	28	93.3	22	23.9	15	21.1

and is relatively easy and inexpensive to record both in the lab and the field without the need for much equipment or 3D scans.

Novices in both groups batter cores at a higher average frequency and with a significantly larger degree of variability than our expert. The observation-only group exhibits the largest variance with some cores bearing no battering and others with striking platforms nearly covered with it. Actively taught novices show a lower average frequency of battering and a slightly lower range of variability than observation only, but still significantly higher than the expert. The expert is far less variable in edge battering with the majority of their cores bearing little to no battering. While this attests to our groups' differing abilities to successfully produce flakes via controlled application of force, it is also likely related to a knapper knowing when a flaking surface or even an entire core is no longer flakable. This can be thought of as knowing when to give up on the one hand and, on the other, as relating to a disposition toward tenacity/determination (Pargeter et al., 2019) and is likely tied with discard behaviors. Novices seem to batter through obstacles when they lack the skill to overcome or avoid them. Our flake invasiveness and edge battering variables are particularly useful for tracking these patterns in the archaeological record because they are relatively straightforward to record.

Our data match up well with Toth et al. (2006) Gona core data. Across the three comparable attributes (edge battering, flake invasiveness, and remaining cortex), Gona consistently shows intermediate degrees of variation between our expert and actively taught novices. Hominins at Gona show a degree of core battering between our expert and actively taught novices. This suggests that Gona knappers were less variable in their ability to successfully produce flakes without damaging the platform than would be expected if knapping behaviors were carried out at low levels of performance. Toth et al. (2006) data also suggest that knappers at Gona displayed judgment about when to give up on a core comparable to our expert and actively taught novices.

Similarly, Gona hominins show a degree of variance in flake invasiveness intermediate between our expert and actively taught novices and a comparably high degree of performance as our expert. The majority of the Gona cores show invasive flake removals much like our expert attesting to a similar consistency in invasive flaking. Our results show that even modern human learners will show a few hours of highly variable knapping without instruction and this can be much more rapidly reduced through instruction. If we assume that the Gona assemblages represent an entire community of knappers, including children/novices, then the high and invariant skill suggest very rapid learning akin to our instructed participants.

Comparing variance in the remaining cortex between our expert and novices and Gona we find, once again, comparable ranges of variability. Gona hominins are largely comparable to our expert and actively taught novices and much less like our observation-only group. This further suggests the potential role of more direct, active teaching in the Oldowan and the importance of skilled individuals in the successful transmission and maintenance of stone technologies over long periods of time. Our results highlight the importance of direct-active teaching in the efficient acquisition of technical skill and the development of regular and consistent control over the knapping process. Our preliminary comparisons to Gona Oldowan archaeological material support this as the case even in the earliest parts of the archaeological record. This supports transmission and transgenerational upkeep of lithic technologies via direct-active teaching from skilled individuals. Our experimental data compared to the Oldowan archaeological data from Gona do not indicate vast differences in uniformity between early hominin knappers and modern knappers. It should also be noted that the raw materials used at Gona, hard, round river cobbles, are much more difficult to knap than the angular basalt used in this experiment. In fact, the same expert knapper from this study achieved a much lower degree of core reduction on the Gona raw materials (between 50 and 60% on average) compared to the basalt (average of 96%) (see Table 2 in Stout et al., 2019). This leaves the possibility that hominins at Gona would have displayed even higher levels of skill had they had access to the same raw materials used in this experiment.

Snyder et al. (2022) recently argued that early stone knapping does not necessitate social transmission of behavioral form and should thus be attributed to low fidelity social learning mechanisms combined with independent individual innovation. Our study, like that of Morgan et al., 2015, addresses the separate question of potential benefits associated with direct active instruction. While it may be possible to re-discover various Oldowan methods and techniques, this does not mean it was easy or likely for Oldowan hominins (Stout et al., 2019), and we should expect the hypothesized fitness benefits of stone tool making (Shea, 2017) to put pressure on the evolution of a more reliable means of transmission even at the earliest stages of this technology. Further, regarding animal behavior studies, there is no reason to assume that the simplest explanation is most likely correct and higher resolution experimental designs are necessary to address competing theories (Heyes, 2012). Our experiment tracks several variables related to the maintenance and management of Oldowan-like core reduction. Collectively, these variables demonstrate the key differences between expert and novice knapping abilities and the role(s) of skill in managing Oldowan-like reduction sequences.

We know that lithic variability is driven by a number of factors, including raw material properties (Jones, 1994; Proffitt et al., 2022) and possibly species-level differences in biomechanics (Key & Dunmore, 2018) and cognition (Assaf et al., 2016). Skill is an additional potential driver of lithic variability in the Oldowan, which may also underlie or interact with others, such as biomechanics (Williams-Hatala et al., 2020). Given the amount of practice required to master stone tool making, much of the record must contain artefacts and possibly entire assemblages created by novices (Assaf et al., 2016; Herzlinger et al., 2017). This, along with differences in knapping strategies (Stout et al., 2019), is a likely factor behind the wide range of variability observed in the Oldowan across both space and time, even in roughly contemporaneous localities such as Lokalalei 1A and 2C (Delagnes & Roche, 2005; Kibunjia 1994) and Gona.

Although the work presented in this paper would benefit from a larger, more diverse sample and longer training times, the experimental exploration of the link between uniformity and skill in the Early Stone Age is a relatively novel approach. Given the results of previous work on the importance of extended practice time in knapper skill acquisition (Pargeter et al., 2019), we expect the differences between actively taught and observation-only novices to become even more significantly different from one another and actively taught novices to approach the expert's level and range of variation as practice time increased. With a larger sample it would be useful to incorporate a variety of raw materials to investigate the intra-individual effects of raw material properties and further establish archaeologically relevant and compatible results. Further, the current experiment included only one expert knapper/instructor, which inherently creates some bias towards the idiosyncrasies and instructional style of that expert. This likely contributes to some of the variation between our results and other studies including expert stone workers. Additional experiments may benefit from a larger sample of experts as well as comparisons between them to better understand how individual expert variation affects both skill manifestation as well as teaching/demonstration and thus transmission.

Better comparisons to currently known Oldowan material need to be made to further contextualize our results within the archaeological record, which requires compatible datasets. By applying our adapted attribute list to Oldowan archaeological assemblages we can then begin doing these comparisons at a variety of scales both spatial and temporal. This higher resolution may shed light on how variability in the OIC changes, if at all, over time and varies across space. Perhaps as the result of variably skilled stone tool markers within a species or different hominin species endowed with fundamental differences in their toolmaking ability.

Conclusions

This study contributes to a growing body of research on the cognitive and perceptual-motor foundations of stone toolmaking skills by investigating the relative contributions of direct-active teaching and skill differences to core uniformity in Oldowan-like flaking systems. We find that novice knappers produce significantly less uniform core patterning and that direct active instruction with language plays a role in increasing core uniformity within novices. Our findings emphasize the critical role of training, practice, and direct-active teaching in facilitating successful flake production even in relatively simple Oldowan-like flaking systems. Even approximately 2 h of dedicated practice was insufficient for novices in our study to develop these skills. Instead they adopted strategies that reduced flaking invasiveness and core reduction intensity while increasing edge battering frequencies. This included repeated failed attempts to overcome less than ideal platform conditions through force alone resulting in blunted platforms and heavily stepped flaking surfaces that further inhibited their ability to flake efficiently. Actively taught novices displayed some improvement, but even that performance fell significantly short of our expert after ∼2 h of training. Novice cores thus remain well short of both modern expert and Oldowan hominin performance, which would likely have taken many hours to achieve. Placed in a comparative framework with the tool skills of earlier hominins, these relatively intense demands suggest that biocultural mechanisms supporting apprenticeship learning would all have been likely targets of selection acting on Oldowan toolmaking aptitudes.

Acknowledgements

The authors wish to thank Emory University and New York University for internal funding that supported this pilot research.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Evan Patrick Wilson is a Doctoral Candidate at the City University of New York. His research focuses on the early hominin record in Africa and the role of cultural evolution and geoarchaeological processes in this record.

Dietrich Stout is Professor of Anthropology at Emory University. His research focuses on the evolution of brain and cognition in the Paleolithic record.

Cheng Liu is a doctoral candidate in the Department of Anthropology at Emory University. His main research interests include Old World prehistory, hunter-gatherer societies, lithic analysis, and cultural evolution. He has conducted field and laboratory research in China, Israel, and Ethiopia.

Megan Beney Kilgore is a Doctoral Candidate at Emory University where she studies the contexts in which adults and children learn stone knapping skills.

Justin Pargeter is Assistant Professor of Anthropology at New York University and honorary Research Fellow at the University of the Witwatersrand. His research focuses on human biocultural evolution tracked by the relationships between technology, cognition, and environmental change in the archaeological record.

ORCID

Dietrich Stout http://orcid.org/0000-0003-3555-7637
Cheng Liu http://orcid.org/0000-0002-1950-7376
Justin Pargeter http://orcid.org/0000-0001-6536-8568

References

Apel, J. (2001). *Daggers knowledge and power: The social aspects of flint dagger technology in Scandinavia (2350-1500 cal BC)*. Wikströms.

Apel, J. (2008). Knowledge, know-how and raw material – the production of Late Neolithic flint daggers in Scandinavia. *Journal of Archaeological Method and Theory*, *15*(1), 91–111. https://doi.org/10.1007/s10816-007-9044-2

Assaf, E., Barkai, R., & Gopher, A. (2016). Knowledge transmission and apprentice flint-knappers in the Acheulo-Yabrudian: A case study from Qesem Cave, Israel. *Quaternary International*, *398*, 70–85. https://doi.org/10.1016/j.quaint.2015.02.028

Bamforth, D. B., & Finlay, N. (2008). Introduction: Archaeological approaches to lithic production skill and craft learning. *Journal of Archaeological Method and Theory*, *15*(1), 1–27. https://doi.org/10.1007/s10816-007-9043-3

Barsky, D. (2009). An overview of some African and Eurasian Oldowan sites: Evaluation of hominin cognition levels, technological advancement and adaptive skills. In E. Hovers, & D. R. Braun (Eds.), *Interdisciplinary approaches to the Oldowan* (pp. 39–47). Springer.

Bleed, P. (2008). Skill matters. *Journal of Archaeological Method and Theory*, *15*(1), 154–166. https://doi.org/10.1007/s10816-007-9046-0

Braun, D. R., Aldeias, V., Archer, W., Arrowsmith, J. R., Baraki, N., Campisano, C. J., Deino, A. L., DiMaggio, E. N., Dupont-Nivet, G., Engda, B., Feary, D. A., Garello, D. I., Kerfelew, Z., McPherron, S. P., Patterson, D. B., Reeves, J. S., Thompson, J. C., & Reed, K. E. (2019). Earliest known Oldowan artifacts at> 2.58 Ma from Ledi-Geraru, Ethiopia, highlight early technological diversity. *Proceedings of the National Academy of Sciences*, *116*(24), 11712–11717. https://doi.org/10.1073/pnas.1820177116

Braun, D. R., Rogers, M. J., Harris, J. W., & Walker, S. J. (2008). Landscape-scale variation in hominin tool use: Evidence from the developed Oldowan. *Journal of Human Evolution*, *55*(6), 1053–1063. https://doi.org/10.1016/j.jhevol.2008.05.020

Bril, B., Rein, R., Nonaka, T., Wenban-Smith, F., & Dietrich, G. (2010). The role of expertise in tool use: Skill differences in functional action adaptations to task constraints. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 825–839. https://doi.org/10.1037/a0018171

Carbonell, E., Barsky, D., Sala, R., & Celiberti, V. (2016). Structural continuity and technological change in Lower Pleistocene toolkits. *Quaternary International*, *393*, 6–18. https://doi.org/10.1016/j.quaint.2015.04.008

Clark, G. (1969). *World history: A new outline*. Cambridge University Press.

Clarkson, C. (2013). Measuring core reduction using 3D flake scar density: A test case of changing core reduction at Klasies River Mouth, South Africa. *Journal of Archaeological Science*, *40*(12), 4348–4357. https://doi.org/10.1016/j.jas.2013.06.007

Cribari-Neto, F., & Zeileis, A. (2010). Beta regression in R. *Journal of Statistical Software*, *34*(2), 1–24. https://doi.org/10.18637/jss.v034.i02

Delagnes, A., & Roche, H. (2005). Late Pliocene hominid knapping skills: the case of Lokalalei 2C, West Turkana, Kenya.

Journal of Human Evolution, *48*(5), 435–472. https://doi.org/10.1016/j.jhevol.2004.12.005

de la Torre, I. (2004). Omo revisited: Evaluating the technological skills of Pliocene hominids. *Current Anthropology*, *45*(4), 439–465. https://doi.org/10.1086/422079

de la Torre, I. (2011). The early stone age lithic assemblages of Gadeb (Ethiopia) and the developed Oldowan/early Acheulean in East Africa. *Journal of Human Evolution*, *60*(6), 768–812. https://doi.org/10.1016/j.jhevol.2011.01.009

de la Torre, I., & Mora, R. (2005). *Technological strategies in the lower Pleistocene at Olduvai Beds I and II*. University of Liege Press.

de la Torre, I., & Mora, R. (2014). The transition to the Acheulean in East Africa: An assessment of paradigms and evidence from Olduvai Gorge (Tanzania). *Journal of Archaeological Method and Theory*, *21*(4), 781–823. https://doi.org/10.1007/s10816-013-9176-5

de la Torre, I., & Mora, R. (2018). Oldowan technological behaviour at HWK EE (Olduvai Gorge, Tanzania). *Journal of Human Evolution*, *120*, 236–273. https://doi.org/10.1016/j.jhevol.2018.04.001

de Lumley, H. (2022). Oldowayen archaïque et Oldowayen classique de l'Afrique à l'Europe méditerranéenne. *L'Anthropologie*, *126*(1), 102993. https://doi.org/10.1016/j.anthro.2021.102993

de Lumley, H., Barsky, D., & Cauche, D. (2009). Archaic stone industries from East Africa and southern Europe Pre-Oldowan and Oldowan. In K. Schick, & N. Toth (Eds.), *The cutting edge. New approaches to the archaeology of human origins* (pp. 55–91). Stone Age Institute Press.

Dibble, H. L., Schurmans, U. A., Iovita, R. P., & McLaughlin, M. V. (2005). The measurement and interpretation of cortex in lithic assemblages. *American Antiquity*, *70*(3), 545–560. https://doi.org/10.2307/40035313

Eerkens, J. W., & Bettinger, R. L. (2001). Techniques for assessing standardization in artifact assemblages: Can we scale material variability? *American Antiquity*, *66*(3), 493–504. https://doi.org/10.2307/2694247

Eren, M. I., Greenspan, A., & Sampson, C. G. (2008). Are Upper Paleolithic blade cores more productive than Middle Paleolithic discoidal cores? A replication experiment. *Journal of Human Evolution*, *55*(6), 952–961. https://doi.org/10.1016/j.jhevol.2008.07.009

Eren, M. I., Lycett, S. J., & Tomonaga, M. (2020). Underestimating Kanzi? Exploring Kanzi-Oldowan comparisons in light of recent human stone tool replication. *Evolutionary Anthropology: Issues, News, and Reviews*, *29*(6), 310–316. https://doi.org/10.1002/evan.21858

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Finestone, E. M. (2019). *Oldowan tool behaviors through time on the Homa Peninsula, Kenya*. City University of New York Graduate Center.

Flenniken, J. J. (1978). Reevaluation of the Lindenmeier Folsom: A replication experiment in lithic technology. *American Antiquity*, *43*(3), 473–480. https://doi.org/10.2307/279406

Gallotti, R., & Mussi, M. (2015). The unknown Oldowan:∼ 1.7-million-year-old standardized obsidian small tools from Garba IV, Melka Kunture, Ethiopia. *PLoS One*, *10*(12), e0145101. https://doi.org/10.1371/journal.pone.0145101

Gowlett, J. A., Stanistreet, I. G., Albert, R. M., Blackbird, S. J., Herries, A. I., Hoare, S., Kogai, P., Komboh, C. K., Mark, D. F., Muriuki, R. M., Murphy, H., Rucina, S. M., & Stollhofen, H. (2022). New Oldowan localities at high level within Kilombe Caldera, Kenya. *L'Anthropologie*, *126*(1), 102976. https://doi.org/10.1016/j.anthro.2021.102976

Guilmet, G. M. (1977). The evolution of tool-using and tool-making behaviour. *Man*, *12*(1), 33–47. https://doi.org/10.2307/2800992

Harmand, S., Lewis, J. E., Feibel, C. S., Lepre, C. J., Prat, S., Lenoble, A., Boës, X., Quinn, R. L., Brenet, M., Arroyo, A., Taylor, N., Clément, S., Daver, G., Brugal, J.-P., Leakey, L., Mortlock, R. A., Wright, J. D., Lokorodi, S., Kirwa, C., … Roche, H. (2015). 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature*, *521*(7552), 310–315. https://doi.org/10.1038/nature14464

Harush, O., & Grosman, L. (2021). Toward the identification of social signatures in ceramic production – an archaeological case study. *Plos one*, *16*(7), e0254766. https://doi.org/10.1371/journal.pone.0254766

Harush, O., Roux, V., Karasik, A., & Grosman, L. (2020). Social signatures in standardized ceramic production – A 3-D approach to ethnographic data. *Journal of Anthropological Archaeology*, *60*, 101208. https://doi.org/10.1016/j.jaa.2020.101208

Herzlinger, G., Goren-Inbar, N., & Grosman, L. (2017). A new method for 3D geometric morphometric shape analysis: The case study of handaxe knapping skill. *Journal of Archaeological Science: Reports*, *14*, 163–173. https://doi.org/10.1016/j.jasrep.2017.05.013

Heyes, C. (2012). Simple minds: A qualified defence of associative learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1603), 2695–2703. https://doi.org/10.1098/rstb.2012.0217

Ingold, T. (1993). Tool-use, sociality and intelligence. In K. R. Gibson, & T. Ingold (Eds.), *Tools, language and cognition in human evolution* (pp. 429–445). Cambridge University Press.

Isaac, G. L. (1986). Foundation stones: Early artefacts as indicators of activities and abilities. In G. N. Bailey, & P. Callow (Eds.), *Stone Age prehistory: Studies in memory of Charles McBurney* (pp. 221–241). Cambridge University Press.

Jones, P. R. (1994). Results of experimental work in relation to the stone industries of Olduvai Gorge. *Olduvai Gorge*, *5*, 254–298.

Key, A. J., & Dunmore, C. J. (2018). Manual restrictions on Palaeolithic technological behaviours. *PeerJ*, *6*, e5399. https://doi.org/10.7717/peerj.5399

Kibunjia, M. (1994). Pliocene archaeological occurrences in the Lake Turkana basin. *Journal of Human Evolution*, *27*(1-3), 159–171. https://doi.org/10.1006/jhev.1994.1040

Kuhn, S. L. (2010). On standardization in the Paleolithic. In A. Nowell, & I. Davidson (Eds.), *Stone tools and the evolution of human cognition* (pp. 105–134). University Press of Colorado.

Lassen, R., & Williams, T. (2015). Variation in flintknapping skill among Folsom-era projectile point types: A quantitative approach. *Journal of Archaeological Science: Reports*, *4*, 164–173. https://doi.org/10.1016/j.jasrep.2015.09.007

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.

Leakey, M. D. (1971). *Olduvai Gorge: Volume 3, excavations in beds I and II, 1960-1963* (Vol. 3). Cambridge University Press.

Marks, A., Hietala, H. J., & Williams, J. K. (2001). Tool standardization in the Middle and Upper Palaeolithic: A closer look (with comments). *Cambridge Archaeological Journal*, *11*(1), 17–44. https://doi.org/10.1017/S0959774301000026

Monnier, G. F., & McNulty, K. P. (2010). Questioning the link between stone tool standardization and behavioral modernity. In S. Lycett, & P. Chauhan (Eds.), *New perspectives on old stones* (pp. 61–81). Springer, New York.

Morgan, T. J., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., Cross, C. P., Evans, C., Kearney, R., de la Torre, I., Whiten, A., & Laland, K. N. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications*, *6* (1), 1–8. https://doi.org/10.1038/ncomms7029

Nonaka, T., Bril, B., & Rein, R. (2010). How do stone knappers predict and control the outcome of flaking? Implications for understanding early stone tool technology. *Journal of Human Evolution*, *59*(2), 155–167. https://doi.org/10.1016/j.jhevol.2010.04.006

Oakley, K. P. (1969). Man the skilled tool-maker. *Antiquity*, *43* (171), 222–224. https://doi.org/10.1017/S0003598X00107653

O'Brien, M. J., Boulanger, M. T., Buchanan, B., Collard, M., Lyman, R. L., & Darwent, J. (2014). Innovation and cultural transmission in the American Paleolithic: Phylogenetic analysis of eastern Paleoindian projectile-point classes. *Journal of Anthropological Archaeology*, *34*, 100–119. https://doi.org/10.1016/j.jaa.2014.03.001

Olausson, D. J. (2008). Does practice make perfect? Craft expertise as a factor in aggrandizer strategies. *Journal of Archaeological Method and Theory*, *15*(1), 28–50. https://doi.org/10.1007/s10816-007-9049-x

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. https://doi.org/10.1080/00934690.2017.1391649

Pargeter, J., Khreisheh, N., & Stout, D. (2019). Understanding stone tool-making skill acquisition: Experimental methods and evolutionary implications. *Journal of Human Evolution*, *133*, 146–166. https://doi.org/10.1016/j.jhevol.2019.05.010

Pargeter, J., Kreisheh, N., Shea, J. J., & Stout, D. (2020). Knowledge vs. Know-how? Dissecting the foundations of stone knapping skill. *Journal of Human Evolution*, *145*, 102807. https://doi.org/10.1016/j.jhevol.2020.102807

Pargeter, J., Liu, C., Kilgore, M. B., Majoe, A., & Stout, D. (2022). Testing the effect of learning conditions and individual motor/cognitive differences on knapping skill acquisition. *Journal of Archaeological Method and Theory*, 1–45. https://doi.org/10.1007/s10816-022-09592-4

Pelegrin, J. (1990). Prehistoric lithic technology: Some aspects of research. *Archaeological Review from Cambridge*, *9*(1), 116–125.

Pelegrin, J. (1993). A framework for analysing prehistoric stone tool manufacture and a tentative application to some early stone industries. In A. Berthelet, & J. Chavaillon (Eds.), *The use of tools by human and non-human primates* (pp. 302–317). Oxford University Press.

Plummer, T. (2004). Flaked stones and old bones: Biological and cultural evolution at the Dawn of technology. *American Journal of Physical Anthropology*, *125*(S39), 118–164. https://doi.org/10.1002/ajpa.20157

Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance*. Sage Publications, Inc.

Proffitt, T. (2018). Is there a developed Oldowan A at Olduvai Gorge? A diachronic analysis of the Oldowan in bed I and lower-middle bed II at Olduvai Gorge, Tanzania. *Journal of Human Evolution*, *120*, 92–113. https://doi.org/10.1016/j.jhevol.2018.01.006

Proffitt, T., Bargalló, A., & de la Torre, I. (2022). The effect of raw material on the identification of knapping skill: A case study from Olduvai Gorge. *Tanzania. Journal of Archaeological Method and Theory*, *29*, 50–82. https://doi.org/10.1007/s10816-021-09511-z

R Core Team, R. (2013). *R: A language and environment for statistical computing*. Vienna, Austria. https://www.R-project.org/.

Roche, H., Delagnes, A., Brugal, J. P., Feibel, C., Kibunjia, M., Mourre, V., & Texier, P. J. (1999). Early hominid stone tool production and technical skill 2.34 Myr ago in West Turkana, Kenya. *Nature*, *399*(6731), 57–60. https://doi.org/10.1038/19959

Roux, V., & David, E. (2005). Planning abilities as a dynamic perceptual-motor skill: An actualist study of different levels of expertise involved in stone knapping. In V. Roux, & B. Bril (Eds.), *Stone knapping: The necessary conditions for a uniquely hominin behaviour* (pp. 91–116). McDonald Institute for Archaeological Research.

Sahnouni, M., Hadjouis, D., Van Der Made, J., Canals, A., Medig, M., Belahrech, H., Harichane, Z., & Rabhi, M. (2002). Further research at the Oldowan site of Ain Hanech, North-eastern Algeria. *Journal of Human Evolution*, *43*(6), 925–937. https://doi.org/10.1006/jhev.2002.0608

Sánchez-Yustos, P. (2021). Knocking on Acheulean's door. DK revisited (Bed I, Olduvai, Tanzania). *Journal of Archaeological Science: Reports*, *35*, 102763. https://doi.org/10.1016/j.jasrep.2020.102763

Sánchez-Yustos, P., Diez-Martín, F., Domínguez-Rodrigo, M., Duque, J., Fraile, C., Baquedano, E., & Mabulla, A. (2017). Diversity and significance of core preparation in the Developed Oldowan technology: Reconstructing the flaking processes at SHK and BK (middle-upper Bed II, Olduvai Gorge, Tanzania). *Boreas*, *46*(4), 874–893. https://doi.org/10.1111/bor.12237

Semaw, S., Renne, P., Harris, J. W., Feibel, C. S., Bernor, R. L., Fesseha, N., & Mowbray, K. (1997). 2.5-million-year-old stone tools from Gona, Ethiopia. *Nature*, *385*(6614), 333–336. https://doi.org/10.1038/385333a0

Semaw, S., Rogers, M. J., Quade, J., Renne, P. R., Butler, R. F., Dominguez-Rodrigo, M., Stout, D., Hart, W. S., Pickering, T., & Simpson, S. W. (2003). 2.6-Million-year-old stone tools and associated bones from OGS-6 and OGS-7, Gona, Afar, Ethiopia. *Journal of Human Evolution*, *45*(2), 169–177. https://doi.org/10.1016/S0047-2484(03)00093-9

Shea, J. J. (2017). Occasional, obligatory, and habitual stone tool use in hominin evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *26*(5), 200–217. https://doi.org/10.1002/evan.21547

Snyder, W. D., Reeves, J. S., & Tennie, C. (2022). Early knapping techniques do not necessitate cultural transmission. *Science Advances*, *8*(27), eabo2894. https://doi.org/10.1126/sciadv.abo2894

Stiles, D. (1979). Early acheulian and developed oldowan. *Current Anthropology*, *20*(1), 126–129. https://doi.org/10.1086/202218

Stout, D. (2002). Skill and cognition in stone tool production: An ethnographic case study from Irian Jaya. *Current Anthropology*, *43*(5), 693–722. https://doi.org/10.1086/342638

Stout, D. (2005). The social and cultural context of stone-knapping skill acquisition. In V. Roux, & B. Bril (Eds.), *Stone knapping: The necessary conditions for a uniquely hominin behaviour* (pp. 331–340). McDonald Institute for Archaeological Research.

Stout, D., Rogers, M. J., Jaeggi, A. V., & Semaw, S. (2019). Archaeology and the origins of human cumulative culture: A case study from the earliest Oldowan at Gona, Ethiopia. *Current Anthropology*, *60*(3), 309–340. https://doi.org/10.1086/703173

Stout, D., Schick, K., & Toth, N. (2009). Understanding Oldowan knapping skill: An experimental study of skill acquisition in modern humans. In K. Schick, & N. Toth (Eds.), *The cutting edge: New approaches to the archaeology of human origins* (pp. 247–266). Stone Age Institute Press.

Stout, D., Semaw, S., Rogers, M. J., & Cauche, D. (2010). Technological variation in the earliest Oldowan from Gona, Afar, Ethiopia. *Journal of Human Evolution*, *58*(6), 474–491. https://doi.org/10.1016/j.jhevol.2010.02.005

Torres, C., & Preysler, J. B. (2020). Experts also fail: A new methodological approach to skills analysis in lithic industries. *Journal of Paleolithic Archaeology*, *3*(4), 889–917. https://doi.org/10.1007/s41982-020-00063-4

Toth, N., & Schick, K. (2018). An overview of the cognitive implications of the Oldowan Industrial Complex. *Azania: Archaeological Research in Africa*, *53*(1), 3–39. https://doi.org/10.1080/0067270X.2018.1439558

Toth, N., Schick, K., & Semaw, S. (2006). A comparative study of the stone tool-making skills of Pan, Australopithecus, and Homo sapiens. In K. Schick, & N. Toth (Eds.), *The Oldowan: Case studies into the earliest stone age* (pp. 155–222). Stone Age Institute Press.

Trigger, B. (1996). *A history of archaeological thought* (2nd ed.). Cambridge University Press.

Wendrich, W. (Ed.). (2013). *Archaeology and apprenticeship: Body knowledge, identity, and communities of practice*. University of Arizona Press.

Williams-Hatala, E. M., Hatala, K. G., Key, A., Dunmore, C. J., Kasper, M., Gordon, M., & Kivell, T. L. (2020). Kinetics of stone tool production among novice and expert tool makers. *American Journal of Physical Anthropology*, *174*(4), 714–727. https://doi.org/10.1002/ajpa.24159