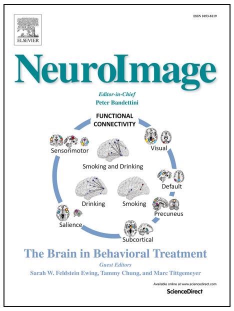
Accepted Manuscript



Prefrontal cortex activation supports the emergence of early stone age toolmaking skill

Shelby S.J. Putt, Sobanawartiny Wijeakumar, John P. Spencer

PII: S1053-8119(19)30448-3

DOI: https://doi.org/10.1016/j.neuroimage.2019.05.056

Reference: YNIMG 15895

To appear in: *NeuroImage*

Received Date: 30 January 2019

Revised Date: 20 May 2019

Accepted Date: 21 May 2019

Please cite this article as: Putt, S.S.J., Wijeakumar, S., Spencer, J.P., Prefrontal cortex activation supports the emergence of early stone age toolmaking skill, *NeuroImage* (2019), doi: https://doi.org/10.1016/j.neuroimage.2019.05.056.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Prefrontal cortex activation supports the emergence of Early Stone Age toolmaking skill

Authors: Shelby S. J. Putt*[a,b,1], Sobanawartiny Wijeakumar[c], John P. Spencer*[d]

[a]The Stone Age Institute, 1392 West Dittemore Road, Gosport, IN 47433, USA.
[b]Center for Research into the Anthropological Foundations of Technology, Indiana University, Bloomington, IN, USA. Email: ssputt@indiana.edu

[c]Faculty of Natural Sciences, University of Stirling, Stirling, Scotland, FK9 4LA, UK. Email: Sobanawartiny.wijeakumar@stir.ac.uk

[d]School of Psychology, University of East Anglia, Lawrence Stenhouse Building 0.09, Norwich Research Park, Norwich, Norfolk, NR4 7TJ, UK. Email: J.Spencer@uea.ac.uk

[1]Permanent address: Department of Sociology & Anthropology, Campus Box 4660, Schroeder Hall 332, Illinois State University, Normal, IL, 61790, USA

*Corresponding authors: ssputt@indiana.edu, J.Spencer@uea.ac.uk

Declarations of interest: none

ACCEPTED MANUSCRIPT

Abstract

Trends toward encephalization and technological complexity ~1.8 million years ago may signify cognitive development in the genus *Homo*. Using functional near-infrared spectroscopy, we measured relative brain activity of 33 human subjects at three different points as they learned to make replicative Oldowan and Acheulian Early Stone Age tools. Here we show that the more complex early Acheulian industry recruits left dorsolateral prefrontal cortex when skills related to this task are first being learned. Individuals with increased activity in this area are the most proficient at the Acheulian task. The Oldowan task, on the other hand, transitions to automatic processing in less than four hours of training. Individuals with increased sensorimotor activity demonstrate the most skill at this task. We argue that enhanced working memory abilities received positive selection in response to technological needs during the early Pleistocene, setting *Homo* on the path to becoming human.

Key Words

motor learning, cognitive evolution, stone tools, fNIRS, neuroarchaeology, working memory

1. Introduction

One cognitive domain in which *Homo sapiens* appear to depart from the great ape pattern is executive functioning. Executive functions include a variety of cognitive processes that allow one to mentally manipulate information, think before acting, solve novel problems, resist temptations, and focus attention (Diamond, 2013). For example, humans tend to outperform other great apes in working memory (WM) tasks (Barth & Call, 2006; Washburn et al., 2007). WM is a system that activates and sustains a set of mental representations for further manipulation and processing (Carruthers, 2013). While WM does not reside in a single neural structure, the dorsolateral prefrontal cortex (dlPFC), with its many cortical and subcortical connections, is thought to play a major role in WM functions in humans (Barbey et al., 2013) and in other primates (Carruthers, 2013; Fuster, 2000; Goldman-Rakic, 1995; Petrides, 2000). It is therefore unsurprising that the dlPFC is one of several areas in the cerebral cortex that has expanded relative to other areas of the brain over the course of human evolution (Van Essen & Dierker, 2007).

This trend toward encephalization in early humans was likely accompanied by or possibly even caused by the enhancement of cognitive features like WM (Sherwood et al., 2008). The first major encephalization event occurred between 1.6 and 1.8 million years ago (Ma) (Shultz et al., 2012) and coincides with the emergence of *Homo erectus sensu lato* and the appearance of a relatively more complex stone tool industry called the early Acheulian. Although earlier accounts of human cognitive evolution have remarked upon the limited WM capacity of early *Homo* (Coolidge & Wynn, 2001, 2005; Wynn & Coolidge, 2004), more recently, there has been accumulating evidence suggestive of an enhancement in executive functions and component cognitive processes associated with the Acheulian industry (Coolidge & Wynn, 2016; Henshilwood & Dubreuil, 2011; Putt et al., 2017; Read, 2008; Stout et al., 2014; 2015; Wynn & Coolidge, 2016).

ACCEPTED MANUSCRIPT

As products of cognition in action, archaeological artefacts can be used to test this hypothesis using a neuro-archaeological approach. Specifically, neuroimaging methods can be combined with experimental archaeology to probe the functional neural processes that underlie tool production, making it possible to identify the cognitive features that past hominins may have used to make certain types of stone tools. The results of a recent study using this approach suggest that early Acheulian tool production, when contrasted with simpler Oldowan toolmaking, relies on a WM network to coordinate between the different goals of the task (Putt et al., 2017). Critically, however, dlPFC activation was absent in this and previous studies that measure brain activity related to stone tool production (Stout & Chaminade, 2007; Stout et al., 2008). An fMRI study found left mid-dlPFC activation when modern subjects trained in Oldowan and Acheulian toolmaking methods made technical judgments about planned actions on partially completed Acheulian tools (Stout et al., 2015), but whether stone tool manufacture elicits dlPFC activation and associated WM functions or not remains an open question.

It is possible that dlPFC activation is present during stone tool manufacture but went unnoticed in our previous study because of the type of contrast analysis employed (Oldowan vs. Acheulian). A follow-up region-of-interest (ROI) study found that bilateral dlPFC is significantly activated during early Acheulian toolmaking relative to a resting state, but only to a limited extent (Putt & Wijeakumar, 2018). Another possibility that we probe here is that dlPFC activation went unnoticed in our previous report because we measured brain activity after participants had completed the training program. Decreased activation in dlPFC is often reported after training on a complex task that involves WM (Jansma et al., 2001; van Raalten et al., 2008). This is consistent with other studies showing that a more precise functional map with fewer activated areas emerges over learning as neural processing becomes more efficient (Garavan et al., 2000; Landau et al., 2004). We therefore expect that the functional neuroanatomy of Oldowan and Acheulian tool production is different at early stages of learning because of differences in task complexity. If this is the case, the increased WM demands of the Acheulian task may elicit increased activation of dlPFC *during earlier stages in training*. This would imply that WM was a cognitive strategy used by early *Homo* to acquire the skills related to making an Acheulian handaxe.

To test this hypothesis, we trained participants to make stone tools that attempt to replicate Oldowan and early Acheulian industries from the archaeological record. At three separate points in the participants' training, we measured real-time changes in oxygenated and deoxygenated haemoglobin (HbO and HbR, respectively) in the cortex using image-based functional near-infrared spectroscopy (fNIRS). This approach produces reconstructed images of localized functional brain activity that can be directly compared to fMRI results (Wijeakumar et al., 2015, 2017). During these neuroimaging sessions, participants engaged in Oldowan and early Acheulian toolmaking tasks, and we assessed how cognitive networks associated with learning to make early forms of stone technology change with training.

2. Methods

2.1. Participants

This study included 33 healthy, right-handed, adult participants (17 females, 16 males; age 23.8 ± 7.9 years [mean ± SD) who had no previous stone knapping experience. The Benton Neuropsychology Clinic Handedness test was administered during the screening process to determine the laterality quotient of potential subjects (Oldfield, 1971). Only subjects who fell within the range of +75 to +100 points (i.e., extreme right-handedness) were included in the experiment. The study was approved by the IRB and Human Subjects Office at the University of Iowa (IRB ID #: 201304789), and all research was performed in accordance with IRB and human subjects guidelines. All subjects signed an informed consent document prior to participating and were paid for their time during the experiment.

Participants were randomly assigned to one of two social transmission groups, which determined whether they received verbal instructions (n = 17) or no verbal cues (n = 16) while watching video tutorials (see supplementary materials for group results). This was to ensure that any activation of higher-order cognition areas could not be attributed to receiving spoken instructions, a modern learning strategy that may not have been available to early *Homo* toolmakers.

The manual dexterity of each volunteer was measured using the Minnesota Manual Dexterity Test (Yankosec & Howell, 2009). Participants were divided into the two learning groups based on their performance on this test so that dexterity levels were equally distributed across groups. The nonverbal group averaged 182.4 ± 17.5 s to place all sixty pegs in the holes on the board in three iterations, while the verbal group averaged a nearly identical 182.7 ± 16.9 s. There was no significant difference in dexterity between the two groups based on this assignment (t = 0.06, p = 0.95). Also, males, who averaged 181.4 ± 14.2 s, and females, who averaged 183.6 ± 19.5 s, did not significantly differ from each other in their dexterity scores (t = -0.34, p = 0.74).

2.2. Training Procedures

The participants individually attended seven 60-min stone knapping training sessions over a period of 3-4 weeks, during which they learned how to knap two different types of archaic stone tools by watching instructional videos. We chose video instruction rather than in-person instruction to ensure that every subject received the exact same instructions at the same rate and also to control for interactive teaching. The videos featured an expert knapper with over 12 years of experience. His face was not visible in the frame, though his hands, lap, and torso were visible. This prevented the nonverbal group from being exposed to any verbal cues that might be communicated by the face. Both groups watched the same instructional videos; however, the nonverbal group watched a silent version. Each practice session proceeded in the following order: 1) a 10-min instruction video; 2) 20 min of individual practice; 3) the same 10-min instruction video repeated; and 4) 20 min more to practice. Subjects were not able to manipulate the video in any way, for example, by pausing it.

All the debitage (toolmaking debris) created while knapping fell on a large tarpaulin mat. After the participants completed a core or core tool and were ready to move on to another rock, the core/core tool and its corresponding debitage were collected, bagged, and labelled with the rock number and other pertinent information for further analysis.

There was relatively little interaction between the experimenters and the participants during the training sessions, except during the first session when the experimenter ensured that the participant understood proper safety protocol. Participants were also told during the first training session to do their best to infer the goals of each training video that they would watch. Explicit goals, such as recognizing ideal platform angles, proper grip of the hammer stone, flake production, alternate flaking, platform preparation, and biface shaping and trimming were stated

ACCEPTED MANUSCRIPT

by the instructor in the videos when unmuted; however, participants in the nonverbal group had to infer these goals from the instructor's actions alone.

Each practice session introduced a new goal for the volunteers to meet or reviewed and refined skills introduced in previous sessions. The skills learned during practice sessions 1 and 2 were comparable to the skills associated with Oldowan simple tool production. This is a quick and expedient method of obtaining a sharp flake to use as a tool (Toth & Schick, 2018). The skills learned during these first two sessions are also essential skills to master in order to make a biface. The first video instructed the participants on how to recognize ideal striking angles on the raw material and create flakes, while the second video reviewed this skill. Practice sessions 3-7 introduced and reviewed skills involved in the production of the early Acheulian technocomplex, which involves a more efficient removal of flakes and the intentional shaping of a large cutting tool (Stout, 2011). The third practice session video featured alternate flaking around a squared edge. The fourth practice session video introduced core bifaces as the instructor demonstrated biface manufacture at a very slow rate. The fifth instruction video focused on primary thinning of a piece to remove large convexities. The sixth instruction video presented information on how to shape and refine a biface by trimming. Finally, the subjects were presented with an instruction video during the seventh practice session that reviewed the entire process of bifacial reduction.

2.3. Neuroimaging Procedures

In addition to the training sessions, participants attended three 90-min neuroimaging sessions after the first, fourth, and seventh training sessions, during which they were video recorded and their brain activity was measured using fNIRS. They sat in a small room surrounded by black curtains. Set-up involved measuring the participant's head to ensure the proper cap size, and measuring 10-20 landmarks to ensure proper cap placement on the head. Hair was cleared at each optode site. The 10-20 landmarks and positions of the sources and detectors on the head were then digitized using a Polhemus Patriot™ Motion Tracking System (Colchester, VT).

2.3.1. Presentation of Stimuli

Stimuli were presented using a block design. The experimental program was designed with EPrime software. Each imaging session consisted of 1) a motor baseline task, 2) an Oldowan toolmaking task, and 3) an Acheulian toolmaking task. The purpose of the motor baseline task was to observe activation of motor-related brain areas while striking rocks together in three different actions that resemble stone knapping (direct percussion, glancing percussion, and grinding) without the added element of actual flake removal. These actions were synchronized to the beat of an externally heard 60 beats per minute metronome initially (10 s). After the metronome concluded, subjects continued the actions to an internal beat that was meant to match the external beat they had just heard (20 s). This task was made up of 9 40-s blocks of activity segregated by 20-s rest periods. The Oldowan toolmaking task was segregated into five 1-min blocks of activity with 15-s resting periods in between each block. The Acheulian toolmaking task was segregated into fifteen 1-min blocks, separated by 15-s rest periods. The order of the tasks was not randomized during each imaging session nor was the length of resting periods; thus, there is some possibility that habituation effects impacted our results. These limitations should be addressed in future studies.

To eliminate the possibility of linguistic contamination, the experiment was designed so that all instructions were given via a silent video, with timing of events indicated by different tones, and subjects were instructed to not talk during the experiment. They were told at the beginning of each neuroimaging session to perform the same activity that they viewed in the instruction videos, which preceded each new task or event. The two instruction videos that preceded the Oldowan and Acheulian tasks were muted and short (~20 s long), featuring the same instructor as the training session videos. These clips showed the final stages of tool manufacture for both tool types so that it was clear to the participants what type of tool they should attempt to make. Instructions also included training on the meanings of different tones that they would hear throughout the session that would signal whether to stop or start an action.

For all training and neuroimaging sessions, subjects were required to wear safety goggles, leather work gloves, and lap pads. They were also given the choice to wear a facemask to block out small particles of airborne silicates.

2.3.2. Materials

At each training and neuroimaging session, subjects were presented with three or four local, granitic rocks of varying sizes that were naturally rounded for use as hammer stones, as well as three siliceous rocks to use as blanks for flake removal. A goal of the training was to introduce the subjects to different qualities, shapes, and types of rock to fracture so that they would learn to select the blank of highest quality and the most workable edges from the three choices that they were always provided. Thus, a variety of unheated cherts from the Midwestern United States, Texas, and California were obtained from collectors in Missouri and Texas, though most of the material was Burlington chert.

Prior to being made available for the subjects to knap, each stone was assigned a unique, identifying label, weighed on a digital scale, and assigned a measurement of volume by the water displacement method. Spalls and cobbles ranged between 69.6 and 3000.0 g in mass (mean = 676.8 g) and had a volume between 20 and 1200 cm^3 (mean = 284.3 cm^3). Generally, smaller pre-made spalls of chert with edges of very acute angles were provided in the first two training sessions. By the third and fourth training sessions, the participants chose from medium-sized spalls without cortex that had edges with more difficult angles, as well as rounded cobbles with cortex but with one or more flakes already removed to help them get started. A mix of small- to medium-sized spalls and cobbles were available to choose for the Oldowan task during the neuroimaging sessions. Larger pieces, some with square edges, were provided for the fifth, sixth, and seventh training sessions and the Acheulian task during the neuroimaging sessions.

2.3.3. Behavioural Data Acquisition

All core and debitage pieces were collected after the completion of each finished core/core tool during the neuroimaging sessions for further analysis. Any debitage that passed through a 6.35 mm screen was discarded. A sample of 17,365 debitage pieces from 235 rocks reduced by all 33 participants in the study was collected and measured for the behavioural analysis. Each piece was weighed to the nearest tenth of a gram and allocated to a metric size category continuum as defined by the smallest of a series of nested squares on centimetre graph paper into which the piece would completely fit (i.e., 1 cm^2, 2 cm^2, 3 cm^2, etc.). All non-core debitage was coded as a flake (either complete, proximal, or distal) or nonflake debitage shatter

(Andrefsky, 2005). Digital callipers were used to measure the maximum thickness for each piece, as well as the maximum platform width and thickness of any flakes with an intact striking platform.

Relative knapping skill was measured using the following variables. The first set of variables measured correspond to flake and platform shape. Platform shape, determined by the ratio of maximum platform width to platform thickness, is a common method used to measure knapping skill (Putt et al., 2014; Stout et al., 2014; Toth et al., 2006), as platform shape contributes to the size and shape of the overall flake. The ratio of flake size to flake mass was also included to determine flake shape differences (Putt et al., 2014; Toth et al., 2006). A larger ratio in both cases signifies a flake that is both relatively thin and elongated, which supposedly demonstrates the knapper’s ability to remove desired flake tools in the case of the Oldowan task and long, thinning flakes for shaping the core tool in the case of the Acheulian task. We calculated the relative platform area ([platform width*platform thickness]/flake size) with the expectation that knappers of a higher skill level would produce smaller, thinner platforms relative to the size of the rest of the flake (Stout et al., 2014).

The second set of variables measured correspond to the efficient use of raw material, as inefficient use of raw material is indicative of low skill level, especially when making Oldowan tools (Bamforth & Finlay, 2008). We examined the proportion of intended flakes to unintended shatter fragments, both on low-quality and high-quality material (Putt et al., 2014; Toth et al., 2006), with the expectation that the assemblages of relatively more skilled knappers would include a higher percentage of flakes than the assemblages of less skilled knappers, demonstrating better control of the material. We also examined the proportion of whole flakes to flake fragments. Previous experimental research demonstrated that the assemblages of skilled knappers included more flake fragments than the assemblages of less skilled knappers, perhaps a combination of skilled knappers striking the core at a higher velocity while attempting to produce thinner, more delicate flakes (Toth et al., 2006). A clear sign of knapping skill is the level of reduction of the cobble into usable flakes (Toth et al., 2006). We measured this by determining the proportion of the original cobble’s mass into flake, shatter, and unexploited core mass, with the expectation that the more skilled knappers would have a larger percentage of flake mass during both Oldowan and Acheulian tasks and a smaller percentage of unexploited core mass during the Oldowan task. We would not expect skilled knappers to exploit most of the core mass while making shaped Acheulian tools, however.

2.3.4. Hemodynamic Signal Extraction

fNIRS data were acquired at 25 Hz with a TechEn CW6 system with wavelengths of 690 nm and 830 nm. Light was delivered to a customized cap via fibre optic cables. Prior to the study, a custom optode geometry was designed to probe ROIs in frontal, temporal, and parietal cortex (see Putt et al., 2017 for more information). The optode geometry included 12 sources and 24 detectors, creating 36 channels with a source-detector separation of 3 cm and two short source-detector channels with a separation of 1 cm. The presentation of stimuli was synchronized with the CW6 system.

HOMER2 software was employed to demean and convert the data into optical density (OD) units. A targeted principal component analysis (tPCA) was applied to data from the three tasks to eliminate noise and motion artefacts (Yücel et al., 2014). We used a general linear model (GLM) to obtain beta values (β) from our block design for HbO and HbR measures in every

ACCEPTED MANUSCRIPT

channel for all conditions in every task for each subject. Signals from short source-detector pairs (channels with the greatest correlation) were used to regress out the effects of superficial layers of the head from signals from the rest of the channels (Gagnon et al., 2011).

2.4. Image Reconstruction

fNIRS data are acquired via sensors placed on the surface of the head. In particular, an optical source is placed near a detector forming a channel. Many studies using fNIRS report data and run statistics on data from each channel. For significant channels, the researchers then infer the locus of brain activity by approximating where the signal is likely to originate from. This can be done by approximating the location of the channel using the 10-20 system of electrode placement or by overlaying the channel on a head model and finding the cortical sites directly below the channel.

Although these approaches are commonly used, they have key limitations. First, it is difficult to place a NIRS cap in exactly the same place from session to session, even on the same individual. This means that there could be some session-to-session variability in the scalp location at which recordings were being performed. Clearly, this is a fundamental limitation if one wants to infer changes in brain activity over learning as in the present study. Second, head sizes differ across individuals; thus, how the channels are laid out on the head of one individual will be different for another individual, particularly if the source-detector distance is held constant.

An alternative to ignoring these sources of spatial variance is to account for them explicitly using image reconstruction techniques. That is, fNIRS data can be moved from channel-space on the surface of the head to voxel-space within the brain volume. This image reconstruction process has been applied successfully in several independent studies by different labs (Eggebrecht et al., 2014; Perlman et al, 2016; Wijeakumar et al., 2015, 2017), and also has been validated by simultaneously measuring brain activity with fNIRS and fMRI (Wijeakumar et al., 2017). Here, we briefly summarize the image reconstruction approach we adopted (see Wijeakumar et al., 2015, 2017 for a more extensive explanation of this process).

Scalp 10-20 landmarks from the session that had the best symmetry were chosen as the reference for each subject. The landmarks from the other two sessions were transformed (linear) to fit this reference set of landmarks. The transformation matrices were applied to the corresponding source and detector positions. AtlasViewerGUI (available within HOMER2) was used to project the points onto an adult atlas using a relaxation algorithm. The projected geometry was used to run Monte Carlo simulations based upon a GPU-dependent Monte Carlo algorithm for each session and subject (Fang & Boas, 2009). This resulted in sensitivity profiles (100 million photons) for each channel of the probe geometry for each session and subject. Head volumes and sensitivity profiles of channels were converted to NIFTI images. Subject-specific head volumes were skull-stripped and transformed to the head volume in the native atlas space using an affine transform (BRAINSFit in Slicer 3D). The transformation matrix obtained was applied to the sensitivity profiles to move them to the transformed head volume space (BRAINSResample in Slicer3D). Sensitivity profiles for all channels were thresholded to include voxels with an OD of greater than 0.0001 (see Wijeakumar et al., 2015 for details). These profiles were summed to create a subject-specific mask for each session, and then these masks were summed across all subjects and sessions (Fig. 1). Only those voxels that contained data from all subjects and all sessions were included in any further analyses. We refer to this image as an intersection mask.

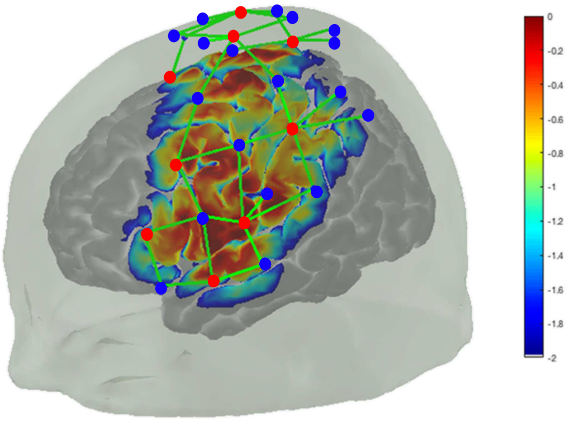


Fig. 1. Probe design with optode positions (red circles represent light sources and blue circles represent light detectors) registered onto an adult atlas head and its corresponding logarithmically-scaled sensitivity map

Image reconstruction combined the beta coefficients for each channel, condition (within each task), and subject with the sensitivity profiles obtained from the Monte Carlo simulations to create voxel-based changes in HbO and HbR concentration (see Wijeakumar et al., 2017 for details). Briefly, the image reconstruction problem can be formulated as the following generic equation:

$$Y = L\,.X \qquad (1)$$

where,

$$Y = \begin{bmatrix} \beta_{dOD}^{\lambda 1} \\ \beta_{dOD}^{\lambda 2} \end{bmatrix}$$

$$L = \begin{bmatrix} \varepsilon_{oxy-Hb}^{\lambda 1}\,.\,F^{\lambda 1} & \varepsilon_{deoxy-Hb}^{\lambda 1}\,.\,F^{\lambda 1} \\ \varepsilon_{oxy-Hb}^{\lambda 2}\,.\,F^{\lambda 2} & \varepsilon_{deoxy-Hb}^{\lambda 2}\,.\,F^{\lambda 2} \end{bmatrix}$$

$$X = \begin{bmatrix} \Delta oxy - Hb_{vox} \\ \Delta deoxy - Hb_{vox} \end{bmatrix}$$

Inverting L to solve for X results in an ill-conditioned and under-determined solution that might be subject to rounding errors. An alternative is to use Tikhonov regularization (Tikhonov, 1963). In this case, the above ‘system’ can be replaced by a regularized ‘system.’ The solution is given by the Gauss-Markov equation,

$$X = (L^T\,L + \lambda.I)^{-1}\,L^T\,.\,Y \qquad (2)$$

where λ is a regularization parameter that determines the amount of regularization and I is the identity operator.

The solution to (2) can be found by minimizing the cost function (Calvetti et al., 2000),

ACCEPTED MANUSCRIPT

$$cost \min X = |L.X - Y|^2 + \lambda . |X - X_o|^2 \qquad (3)$$

where the size of the regularized solution is measured by the norm λ . $|X - X_0|^2$. X_0 is an *a priori* estimate of X, which is set to zero when no priori information is available. Here X is determined for each chromophore and condition separately. Once Equation (3) is solved, there is now a voxel-wise estimate of the concentration data. Thus, the best estimate of the channel-wise concentration data for each condition (from the GLM) has been combined with information from the photon migration results to create an estimate of the voxel-wise concentration data for each chromophore, for each condition, and for each subject.

The resultant beta maps were intersected with the intersection mask to restrict analyses to the voxels that were common to all sessions and subjects. Consequently, voxel-based changes in HbO and HbR concentration were obtained for each condition (within each task) and subject.

2.5. Statistical Analysis

2.5.1. Preliminary Analysis of Motor Baseline Data

Initial examinations of the motor baseline data revealed that performance varied from session to session. Therefore, a Pace (external, internal) x Session (1-3) ANOVA was performed for each of the three conditions of the baseline task (Direct, Glancing, and Grinding) to identify which condition had the fewest number of session-related effects, for the purpose of identifying a stable motor baseline to contrast with the knapping tasks. The ANOVA was conducted with the *3dMVM* function in AFNI (Analysis of Functional Images) (Chen et al., 2014). There was a significant effect of Session for each condition ($F = 3.153$, $p < 0.05$). The glancing condition was selected as the baseline for this study because it most closely resembles the knapping gesture used during the Oldowan and Acheulian tasks, and its combined significant clusters had the fewest number of voxels of the three conditions in the Session effect, meaning this condition remains the most stable over time.

2.5.2. Analysis of Neuroimaging Data

Two separate multi-factorial ANOVA tests were conducted on the HbO and HbR beta maps, with Task (Oldowan, Acheulian) and Session (1-3) as within-subject factors and Group (verbal, nonverbal) as a between-subject factor. Resultant functional images of main effects and interactions were corrected for family-wise errors using the 3dClustSim function (corrected at alpha = 0.05, corresponding to a cluster size threshold of > 27 voxels). We analysed the highest-order effect in each spatially unique cluster; thus, main effect areas that overlapped with areas where an interaction occurred between Task and Session, Group and Task, etc., were interpreted based on the interaction effect. Overall, there were 16 instances of overlapping clusters between effects that were assigned to higher-level effects.

Using the coordinates for the centre of mass of activation for each effect, we extracted the beta values in these areas for the Oldowan and Acheulian tasks, the three sessions, and the verbal and nonverbal groups. In cases of a significant interaction, the averaged beta values of related samples from Task and Session were compared using the Wilcoxon signed-rank test, while the averaged beta values for Group were compared using the Mann-Whitney test. We also compared

ACCEPTED MANUSCRIPT

beta values from the knapping conditions to the motor baseline conditions using the Wilcoxon signed-rank test to identify significant clusters that were unique to stone knapping and not simply general motor regions. The effect size for each cluster was calculated using an eta-squared analysis (Fritz et al., 2012). Only those significant clusters where post-hoc tests determined knapping activation to be significantly higher than motor baseline activation were included in the final results discussed in the main text (see Figs. 4, 6, and 8). Because the motor baseline task did not control for auditory stimulation while clicking rocks together, temporal cortex clusters were also included in the final results, even if the signal in these regions was not significantly higher than the motor baseline signal. The CA_ML_18_MNIA atlas was used to assign labels to the centre of mass of significant clusters with AFNI's 'whereamI' function.

Spheres 8 mm in diameter representing visual WM areas were constructed from published coordinates from a recent meta-analysis (Wijeakumar et al., 2015). Overlap between significant toolmaking clusters and constructed spheres was interpreted as evidence for WM involvement during toolmaking tasks.

2.5.3. Analysis of Behavioural Data

Out of all the behavioural measures tested, only three demonstrate an expected increase in skill over time, all of which reflect the efficient use of raw material. These include the proportion of flake mass removed, the proportion of core mass remaining, and the proportion of flakes produced. For both the Oldowan and Acheulian tasks, we performed a repeated measures ANOVA with session (1-3) and group (verbal, nonverbal) as factors on each of these measures, using SPSS software. We also performed LSD pairwise comparisons to determine whether the means from two sessions (e.g., session 1 vs. session 3) were statistically different. Because the data were not normal, we conducted Spearman's rank correlations to determine the strength of the monotonic relationship between these behavioural measures of knapping skill and the neural signals associated with significant clusters for each task and session.

To determine whether significant neural activity simply reflects differences in the frequency of the knapping behaviour (striking a hammer stone against a core or core tool), we used the total number of debitage elements (complete flakes, flake fragments, and shatter fragments) as a rough proxy for knapping frequency. We then performed Pearson correlations to test the extent of a linear relationship between the total number of debitage elements and the neural activation in significant clusters.

2.6. Data Availability

The datasets generated during the current study are available from the corresponding authors upon reasonable request. This form of data sharing complies with the requirements of the funding bodies and with institutional ethical approval.

3. Results

3.1. Behavioural Results

ACCEPTED MANUSCRIPT

Tools and toolmaking debris produced during the Oldowan and Acheulian tasks during the third neuroimaging session of the experiment, while generally smaller, resemble Oldowan and early Acheulian artefacts discovered at the Gona site in Ethiopia (2.6 Ma) and at the Es2-Lepolosi (1.75-1.4 Ma) and Konso (1.5-1.4 Ma) sites in Tanzania and Ethiopia, respectively (Fig. 2; Tables S1-S2) (Beyene et al., 2013; Diez-Martín et al., 2014; Stout et al., 2010).

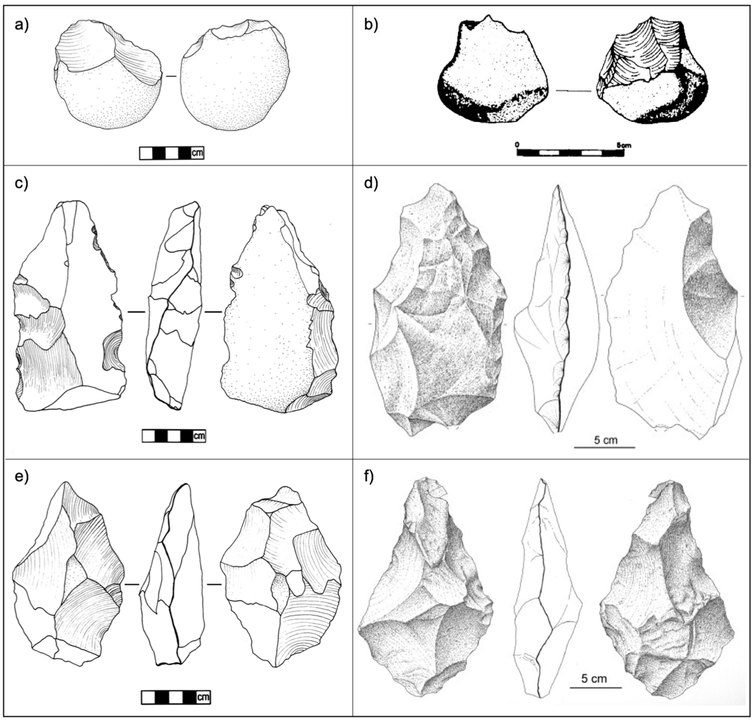


Fig. 2. A comparison of archaeological artefacts to experimental stone tools. A bifacial side chopper produced by one of the participants (a) resembles a unifacial side chopper (EG10) from the early Oldowan site of Gona (reproduced with permission from Springer Nature: *Nature* 2.5-million-year-old stone tools from Gona, Ethiopia, Semaw et al., 1997) (b). Experimental bifaces (c and e) resemble large cutting tools (00/104 and 10/307) from the early Acheulian site of ES2-Lepolosi (d and f) (reproduced with permission from Elsevier: *Quaternary International* Early Acheulean technology at Es2-Lepolosi [ancient MHS-Bayasi] in Peninj [Lake Natron, Tanzania], Diez-Martín et al., 2014). Experimental tool sketches by Jodi Pope Johnson.

The toolmaking debris collected from each participant during the three neuroimaging sessions demonstrates that stone toolmaking skills improved over the course of the experiment (Fig. 3). The participants became more effective at removing flakes from the core over time, as evidenced by a significant increase in the mean percentage of flake mass removed with each session during both the Oldowan task (F = 6.2, p = 0.004; Fig. 3a) and the Acheulian task (F = 3.5, p = 0.037; Fig. 3b). Similarly, they wasted less raw material as they gained proficiency at the Oldowan task, which is reflected by a decrease in the mean percentage of remaining core mass over time (F = 11.8, p < 0.001; Fig. 3c). An increase in the proportion of intentional flakes relative to unintentional pieces of shatter shows that the participants made fewer errors over time during the Oldowan task (F = 2.9, p = 0.066; Fig. 3d). While the increase in the proportion of flakes is not a significant effect overall, an LSD pairwise comparison indicates that there is a significant difference between the first and last session (p = 0.024). Participants in both the

ACCEPTED MANUSCRIPT

verbal and nonverbal learning groups achieved similar levels of proficiency at stone tool manufacture during the experiment (see Fig. S1 and Table S1). In sum, these results demonstrate that greater levels of skill are indicated by an increase in the proportion of flake mass removed for both tasks, while an increase in flakes produced and a decrease in the proportion of core mass remaining is indicative of increased skill for the Oldowan task.

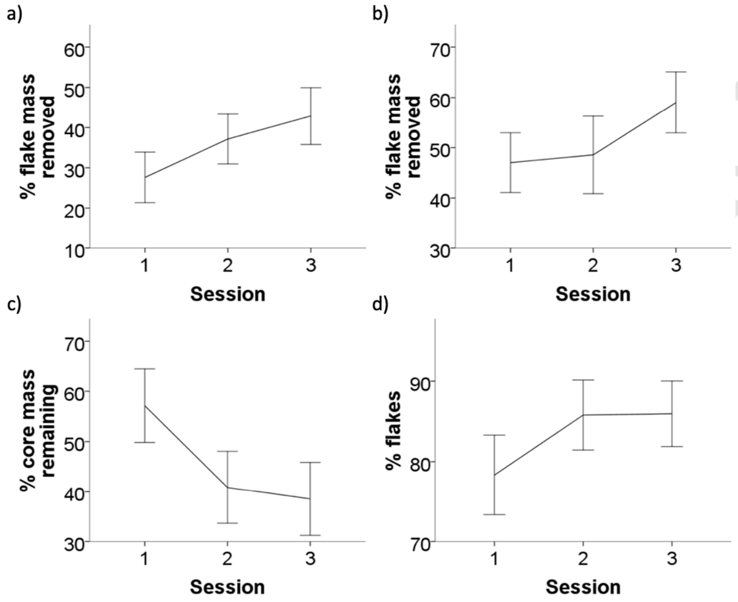


Fig. 3. Behavioural measures that show significant improvement across sessions: mean proportion of flake mass removed from a core relative to total mass of a core prior to reduction during the Oldowan task (a) and Acheulian task (b); mean proportion of core mass remaining after flake removal relative to total mass of a core before reduction during the Oldowan task (c); and the mean proportion of intentional flakes relative to unintentional shatter pieces removed from a core during the Oldowan task (d). Error bars represent 95% confidence intervals.

3.2. fNIRS Results

The main goal of this paper is to test the hypothesis that early Acheulian tool production recruits prefrontal areas that may be involved in WM to a greater extent than Oldowan tool production. Herein, we define HbO neural activity as being “activated” (>0) or “suppressed” (<0) depending on its relationship to the neural state at the start of the block (0) (recall that participants completed a rest phase after each block). As we are primarily interested in effects that generalize across both social transmission groups, this section reports findings from brain areas with changes in HbO that were unaffected by the mode of social transmission assigned to participants (i.e. verbal vs. nonverbal transmission of toolmaking skills). Therefore, this section includes the findings from the Task main effect, Session main effect, and Task x Session effects. All HbO results are reported in Tables 1-3 and S2-S5. To remain consistent with previous studies, the following sections focus only on those clusters that were significantly activated relative to a motor baseline task (marked by asterisks in results tables).

The effect of social transmission, that is, all Group-related effects are reported in Supplementary Materials (Group main effect, Group x Task effect, Group x Session effect, and

ACCEPTED MANUSCRIPT

Group x Task x Session effect). These results demonstrate that the context in which a new motor skill is learned, either with verbal instruction or nonverbal imitation, affects the cognitive strategies used to attend to the task.

HbR results are reported in Supplementary Materials (see Table S6 and Figs. S2-S4). Overall, twelve HbR clusters overlap spatially with significant HbO clusters, and of these, eight show an inverse relationship between HbR and HbO. None of the HbR clusters overlap with HbO clusters that are significantly more active during stone knapping in relation to a simple motor task.

Below, we present the first evidence of increased activity in dlPFC during early Acheulian tool manufacture relative to Oldowan tool manufacture. At no point in the experiment did participants show signs of transitioning to procedural memory while learning to make early Acheulian handaxes like they did while learning to make simple Oldowan flakes. Furthermore, we demonstrate a clear relationship between toolmaking skill and brain activity in different areas of the frontal cortex.

3.2.1. Acheulian vs. Oldowan toolmaking

We identified a total of six clusters where there was a significant difference between the Oldowan and Acheulian toolmaking tasks, that is, a Task main effect (Table 1). Two left hemisphere clusters in the dlPFC, which overlap with the visual WM network (Wijeakumar et al., 2015), exhibited significantly increased neural activity during one of the knapping tasks than during the motor baseline task. These include clusters in the middle frontal gyrus (MFG-1) and the *pars triangularis* in the inferior frontal gyrus (IFG), which extends into MFG. In both cases, activity was greater during the Acheulian task than during the Oldowan task; however, the effect was driven largely by suppression of neural activity during the Oldowan task in the left MFG-1 (Fig. 4a), while in the left IFG, the effect was driven by greater activation during the Acheulian task (Fig. 4b). The MFG-1 cluster overlaps with the precentral gyrus (PrG) cluster that was identified during Acheulian tool production post-training (Putt et al., 2017). Its deactivation during the Oldowan task may be the result of a process called neural repetition suppression, which optimizes the efficiency of neural circuits by facilitating deactivation in dlPFC once learning is successfully completed (León-Carrión et al., 2010).

Table 1. Brain areas showing significant activation in the Oldowan-Acheulian (task) contrast

	Localization[1]	**Sig. Effect**[2]	**MNI Coordinates (mm)**			**Volume (mm^3)**	***M* ΔHbO (μM) ± SEM**	**η^2**
			x	y	z			
Left	Middle frontal gyrus (dlPFC)*	A>O	-47.9	16.1	35.9	5360	6.9 ± 0.07	0.03
Left	Inferior frontal gyrus (dlPFC)*	A>O	-47.2	41.0	12.7	1672	4.90 ± 0.03	0.04
Left	Inferior frontal gyrus	O>A	-58.9	22.1	12.5	480	7.04 ± 0.31	0.04
Right	Paracentral lobule	A>O	7.2	-20.7	82.8	392	5.46 ± 0.12	0.02
Right	Precentral gyrus	A>O	62.9	8.5	29.1	232	5.08 ± 0.13	0.04
Right	Postcentral gyrus	O>A	56.3	-20.7	51.7	232	5.01 ± 0.12	0.02

[1] Areas listed include clusters with a significant task main effect ($p < 0.05$ with family-wise correction using α=0.05) from the Task x Group x Session ANOVA that were not subsumed under a higher-level interaction effect.

[2] A=Acheulian, O=Oldowan

*Indicates cluster where knapping activation is significantly higher than motor baseline activation

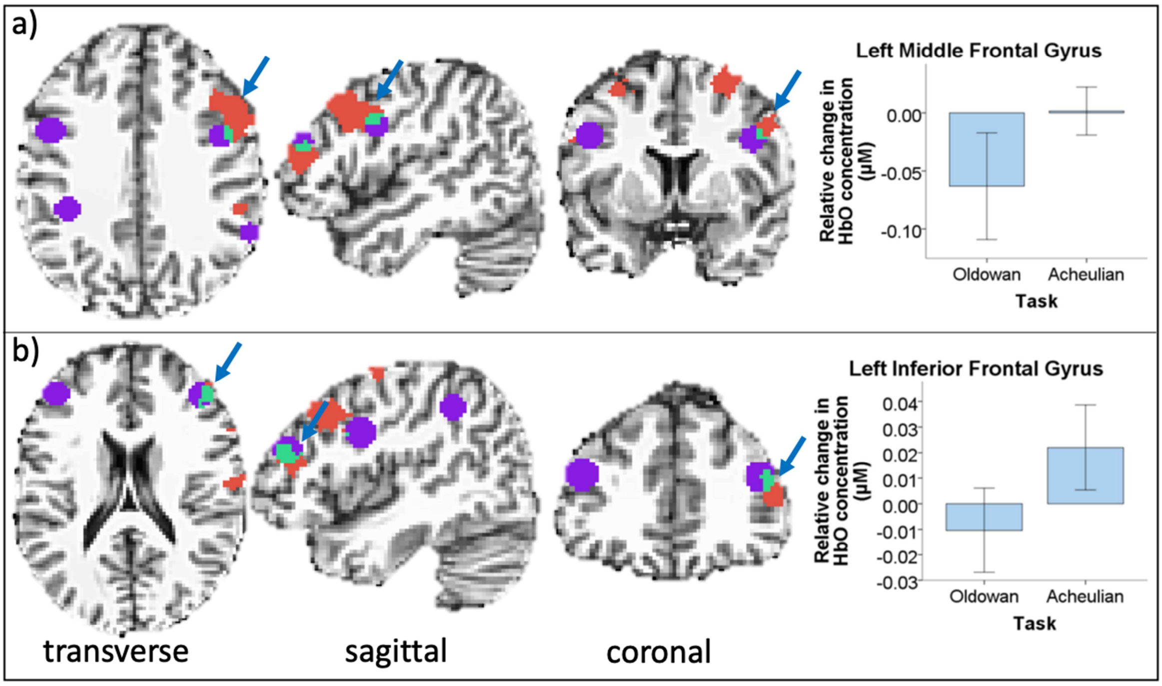


Fig. 4. Significant task results (red) showing greater neural activity in left MFG-1 (a) and IFG (b) clusters in the dlPFC during Acheulian toolmaking compared to Oldowan toolmaking (n = 33, ANOVA F = 4.21, $p < 0.05$). Blue arrows indicate the labelled area. These clusters overlap with visual WM areas (purple). Overlapping voxels are represented by the colour teal. Error bars represent 95% confidence intervals. Figure is in radiological coordinates (left hemisphere is on right side of the transverse and coronal slices).

Although the total number of debitage elements (flakes, flake fragments, and shatter fragments) from a sample of 14,738 was significantly higher during the Acheulian task than the Oldowan task (K-S = 4.0, $p < 0.001$), it was not significantly correlated with activation in either the MFG-1 or IFG clusters (p = 0.126 and 0.100, respectively). Therefore, higher activation of dlPFC during the Acheulian task relative to the Oldowan task cannot be explained simply by a higher frequency of striking the hammer stone against the core tool, the behaviour that this measure roughly estimates. Therefore, a cognitive explanation for the difference in dlPFC activation is warranted.

Moore and colleagues (2006) have shown that increased skill at a visual task is accompanied by increased recruitment of dlPFC during WM encoding and maintenance. We investigated whether a similar pattern of increased dlPFC activation is associated with increased stone toolmaking skill. A positive correlation exists between MFG-1 activation during the Acheulian task and the proportion of flakes produced during the third session (*Spearman's rho* = 0.377, p = 0.040; Fig. 5), showing that the most skilled toolmakers in the study had the highest HbO signal in this WM area of the brain. No such relationship was found between the left IFG cluster and any of the behavioural measures. The correlation between MFG-1 activation and the proportion of flakes produced suggests that the skill required to make flakes during the Acheulian task relies on WM and possibly other cognitive functions such as planning and decision-making.

ACCEPTED MANUSCRIPT

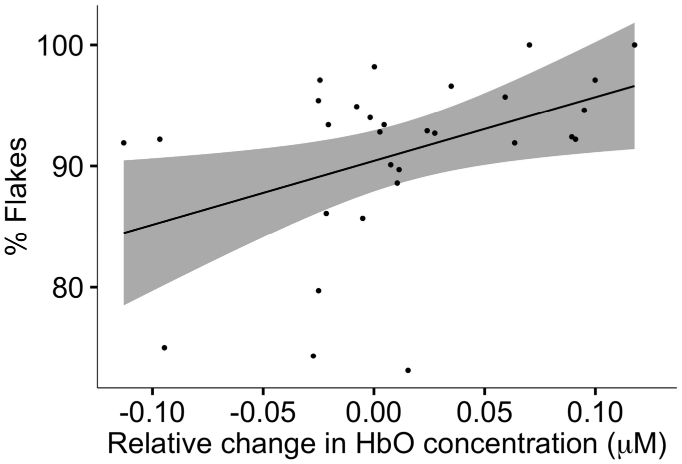


Fig. 5. Positive relationship between left MFG-1 activation during the Acheulian task and the proportion of flakes made by participants in the third session (grey band shows 95% confidence intervals).

3.2.2. Learning networks

Three clusters showed a significant effect across the three sessions for both toolmaking tasks (i.e., a Session main effect), all occurring within the left hemisphere (Table 2). The left dorsal PrG, however, was the only cluster showing significantly greater activity when compared to the motor baseline task, suggesting that the other clusters were mainly involved in improving visuo-motor coordination with training. This result is consistent with studies of motor learning suggesting that the left dorsal PrG contributes to the cognitive aspects of motor learning rather than contributing directly to movement execution (Hardwick et al., 2013). We found that activity in this area decreased from session to session, indicating an increased efficiency with learning/practice (Kelly et al., 2005) (Fig. 6).

Table 2. Brain areas with a significant effect across sessions (both toolmaking tasks included)

Localization[1]		**Sig. Effect**	**MNI Coordinates (mm)**			**Volume (mm^3)**	***M* ΔHbO (μM) ± SEM**	**η²**
			x	y	z			
Left	Superior frontal gyrus	3>2>1	-23.8	3.8	65.5	1528	3.99 ± 0.04	0.01
Left	Precentral gyrus*	1>2>3	-23.5	-18.8	75.3	568	4.05 ± 0.08	0.07
Left	Postcentral gyrus	2>1>3	-60.5	-1.9	26.5	544	3.86 ± 0.06	0.05

[1] Areas listed include session main effect significant clusters ($p < 0.05$ with family-wise correction using α=0.05) from the Task x Group x Session ANOVA that were not subsumed under an interaction effect.

*Indicates cluster where knapping activation is significantly higher than motor baseline activation

The total number of sampled debitage elements increased from a mean of 69.06 during the first session, to 76.47 during the second session, and 85.55 during the third session; however, this change over time was not significant ($F = 1.8$, $p = 0.168$). Moreover, the increase in debitage elements over time was not significantly correlated with activation in dorsal PrG ($p = 0.743$). Therefore, the activation of this cluster does not appear to be associated with the frequency of striking the hammer stone against the core or core tool. Rather, the noticeable decrease and

ACCEPTED MANUSCRIPT

eventual suppression of dorsal PrG activity over the course of the experiment is consistent with dorsal PrG contributing to motor learning.

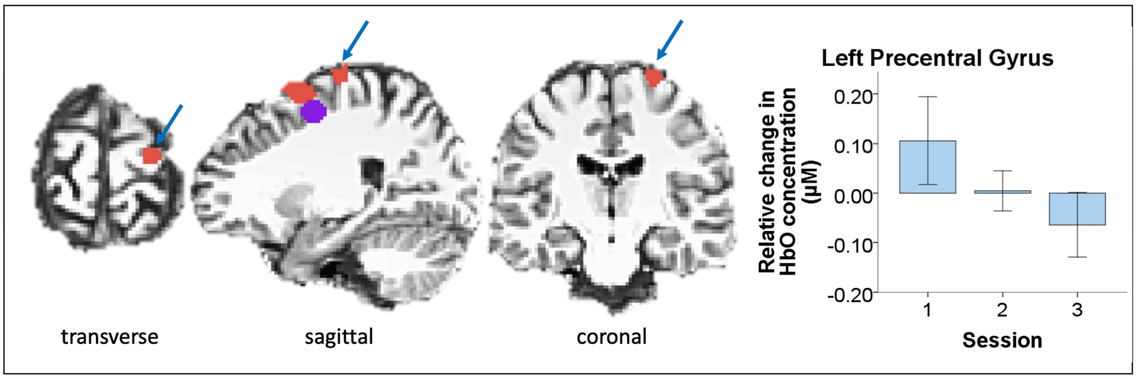


Fig. 6. Significant session result (red), with both toolmaking tasks included, showing a decrease in neuronal activity in dorsal PrG over time as the participants gained more experience in stone knapping ($n = 33$, ANOVA $F = 3.17$, $p < 0.05$). Blue arrows indicate the labelled area. Error bars represent 95% confidence intervals.

Because the dorsal PrG plays a role in learning during the stone knapping tasks, we expected to find a correlation between neural activation in this area and behavioural indices of learning. Deactivation of the left PrG was associated with a greater percentage of flake mass produced in the Acheulian task (*Spearman's rho* = -0.509, $p = 0.004$), which denotes higher skill on the part of the knapper (Fig. 7). Notably, this relationship was statistically robust only by the third session, after participants had the most extensive amount of practice.

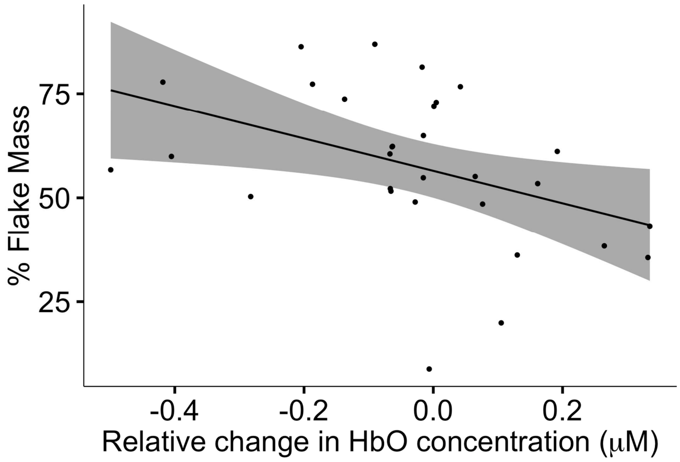


Fig. 7. Significant correlation between participants' left PrG neural activity and the percentage of flake mass during the third session for the Acheulian task (grey band shows 95% confidence intervals).

3.3.3. Learning network differences by task

Eight clusters spanning the frontal, parietal, and temporal cortices were differentially activated depending on the task across the three neuroimaging sessions (i.e., a Session x Task interaction; Table 3). Three of these clusters, including right PrG, postcentral gyrus (PoG), and left MFG-2, exhibited significantly more neural activity during one of the knapping tasks than during the motor baseline task.

The right dorsal PrG cluster showed high activation during the Oldowan task in the first session that decreased in later sessions, while the Acheulian task showed the inverse pattern (Fig. 8a). The PrG cluster falls within or near the frontal eye field (FEF), which is involved in eye movements and associated cognitive processes such as attentional orienting, visual awareness, and decision making, as well as planning complex movements (Vernet et al., 2014). A cluster in the left MFG-2 followed a similar pattern (Fig. 8b).

Table 3. Brain areas that show a significant interaction between task and session

	Localization[1]	**Sig. Effect**[2]	**MNI Coordinates (mm)**			**Volume (mm^3)**	***M* ΔHbO (μM) ± SEM**	η^2
			x	y	z			
Right	Precentral gyrus*	O:1>3>2; A:3>2>1	36.5	-8.1	64.3	3440	4.3 ± 0.04	0.04
Right	Inferior parietal lobule	O:1>3>2; A:3>2>1	55.5	-37.9	46.7	2264	3.9 ± 0.04	0.04
Left	Middle frontal gyrus*	O:1>3>2; A:3>2>1	-37.0	18.5	52.2	1552	3.6 ± 0.02	0.05
Left	Superior temporal gyrus‡	O:2>1>3; A:3>2>1	-58.1	-34.1	18.7	680	3.5 ± 0.03	0.05
Right	Middle temporal gyrus‡	O:1>2>3; A:3>2>1	68.5	-38.6	6.2	400	3.4 ± 0.03	0.04
Left	Inferior frontal gyrus	O:1>3>2; A:2>3>1	-61.5	7.0	15.3	368	3.7 ± 0.09	0.04
Right	Supramarginal gyrus	O:2>1>3; A:2>3>1	64.1	-19.3	36.5	232	3.8 ± 0.10	0.05
Right	Postcentral gyrus*	O:2>3>1; A:1>2>3	37.1	-33.7	69.9	232	4.0 ± 0.10	0.05

[1] Areas listed include session x task interaction effect significant clusters ($p < 0.05$ with family-wise correction using $\alpha=0.05$) from the Task x Group x Session ANOVA that were not subsumed under a larger interaction effect.
[2] A=Acheulian, O=Oldowan
*Indicates cluster where knapping activation is significantly higher than motor baseline activation
‡Temporal areas included in discussion despite not having significantly higher knapping activation than motor baseline activation (see Methods)

Neural activity in the right PoG area increased from the first to second session during the Oldowan task but decreased during the Acheulian task (Fig. 8c). By the third session, there were minimal activation differences between the two tasks. Previous studies looking at brain activation changes across several neuroimaging sessions recorded a similar pattern of decreasing activity in the contralateral primary sensorimotor cortex during the execution of trained hand and wrist movements (Carel et al., 2000; Loubinoux et al., 2001). Together, our results implicate the involvement of this area in a sensorimotor integrative learning process related to the contralateral hand (the left hand in the current study). In the case of Acheulian stone knapping, this might be related to the demands of learning how best to position the core to remove a flake and handling the core with the left hand after delivering a forceful blow with the right hand.

The motor baseline task controlled only for similar motor activity relative to stone knapping and not for the sound of the two rocks when they were struck against each other. For this reason, we would not expect any clusters in the temporal cortex to be significantly more active relative to the motor baseline task. Therefore, we discuss the relevance of two temporal areas that showed a significant interaction effect between task and session. These include the right middle temporal gyrus (MTG) and the left superior temporal gyrus (STG; Fig. 8d-e). Both of these clusters were activated during exclusively Acheulian toolmaking post-training (Putt et al., 2017). The current analysis demonstrates that these temporal areas are likely important for learning during the Acheulian task, as evidenced by an increase in activation over the course of

ACCEPTED MANUSCRIPT

the three sessions. Moreover, similar to our previous findings, right MTG was not heavily recruited at any point in learning during the Oldowan task relative to the Acheulian task, and by the third session, left STG is significantly suppressed during the Oldowan task relative to the Acheulian task and relative to previous sessions.

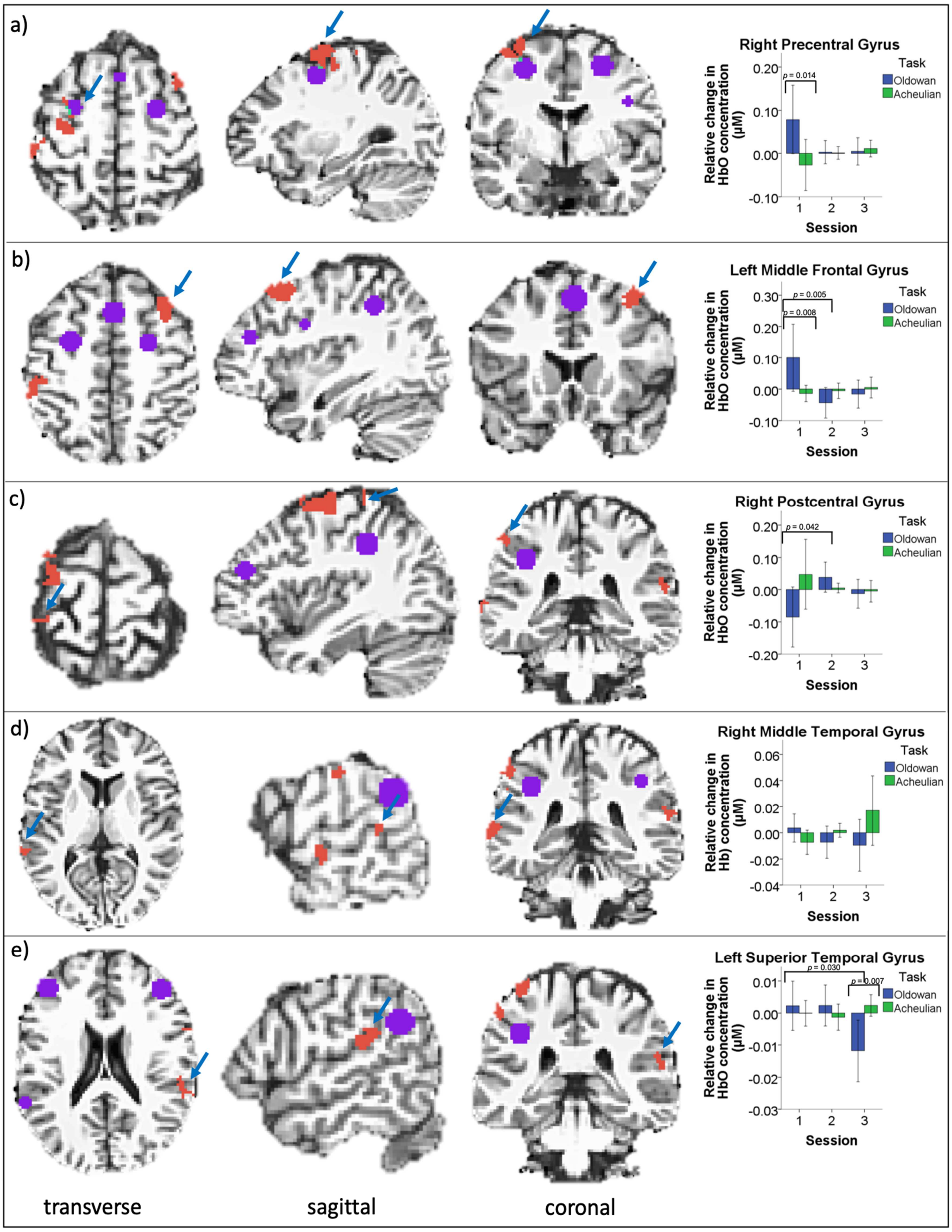


Fig. 8. Significant results (red) showing differential patterns of neural activity between Acheulian and Oldowan toolmaking over the course of three sessions in right PrG in or near to FEF (a), left MFG-2 (b),

right PoG in sensorimotor cortex (c), and right MTG (d) and left STG (e) in the temporal cortex (*n* = 33, ANOVA *F* = 3.17, $p < 0.05$). Blue arrows indicate the labelled area. The right PrG cluster overlaps with a visual working memory area (purple). Overlapping voxels are represented by the colour teal. Error bars represent 95% confidence intervals.

There were significant correlations between neural activity in the right PrG and PoG clusters and the lithic skill measures (Fig. 9). At the group level, the PrG cluster was not strongly associated with Acheulian tool production; however, neural activity in this area was correlated with Acheulian skill acquisition at the individual level (*Spearman's rho* = 0.373, *p* = 0.042 Fig. 9a). Individuals with an activated PrG tended to perform better than others by producing a larger proportion of flakes. This correlation was significant during the third neuroimaging session, at which point there was the highest activation change during the Acheulian task. Because motor learning appears to be associated with relatively decreased activity in the primary sensorimotor cortex by the third session, we expected to find higher skill measures to be negatively associated with neural activity in the right PoG. This was the case in the third neuroimaging session during the Oldowan task: individuals with reduced activity in this area were more likely to leave behind a lower proportion of core mass (*Spearman's rho* = 0.502, *p* = 0.005), which is a sign of enhanced skill in this task.

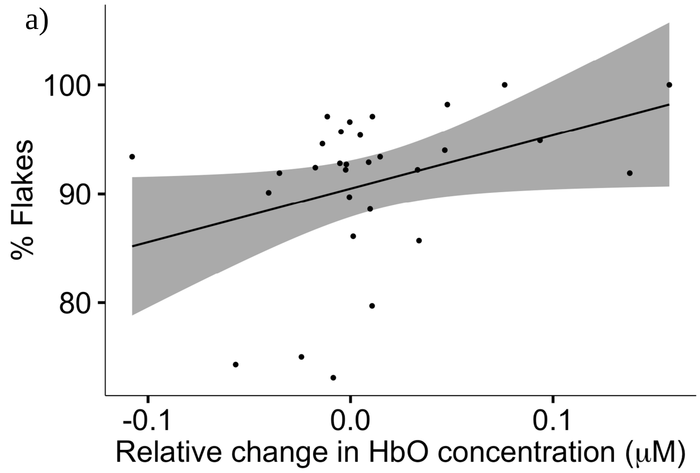
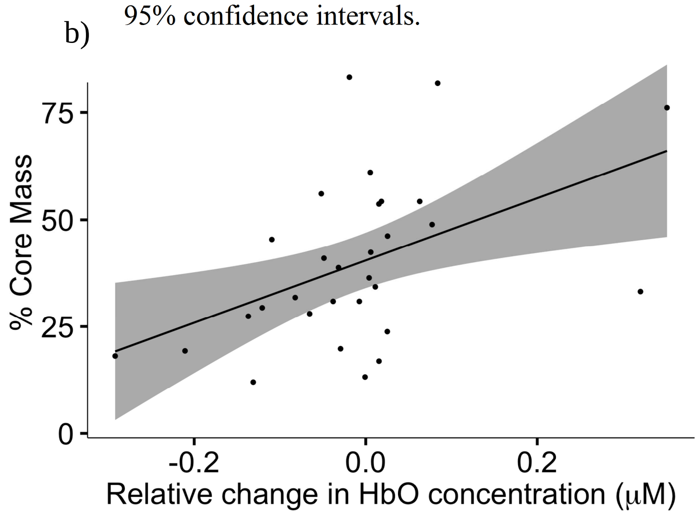


Fig. 9. Relationship between lithic skill measures and neural activity in regions that demonstrate a significant Task x Session interaction effect based on an ANOVA. Neural activity in the right PrG was positively correlated with the proportion of flakes produced during the third session in the Acheulian task (a). Increased skill as measured by the proportion of remaining core mass (b) was associated with higher neural activity in the right PoG during the third session of the Oldowan task. Grey band 95% confidence intervals.



ACCEPTED MANUSCRIPT

4. Discussion and Conclusions

The goal of the present study was to shed new light on questions related to human cognitive evolution based on evidence from functional brain activation data that we collected from modern-day human participants as they replicated naturalistic prehistoric stone tool manufacture. Specifically, we tested whether the complex early Acheulian core-shaping task elicits increased neural activity in the prefrontal cortex relative to the simpler Oldowan flaking task during early stages in training. We also identified the neural sites that are involved in the cognitive aspect of learning these motor skills by comparing data to a motor baseline.

Our main finding is that neural activity in the dlPFC while making early Acheulian tools is significantly higher than it is while making Oldowan tools. WM and potentially other executive functions associated with this area probably play an important role in learning this complex task; however, this remains to be explicitly tested by localizing WM areas within-subjects using an established WM task. Secondly, we found evidence of learning from both lithic debris and neural activation pattern changes across sessions. Furthermore, we show that some lithic indicators of skill may be predictive of frontal activation. This is significant because it could allow archaeologists to infer the level of activation of certain brain areas of once living hominins based on the stone tool artefacts that they left behind, provided that most of the lithic reduction process for a tool is preserved at an archaeological site.

Our results lend support for the hypothesis that selective pressures during the early Pleistocene resulted in an enhancement of WM capacity in early *Homo*. Clusters in left dlPFC are relatively more activated during early Acheulian handaxe production than during Oldowan flake production, especially when the skills related to this task are first being learned. These clusters overlap spatially with the results of a visual WM meta-analysis (Wijeakumar et al., 2015). Furthermore, a significant positive correlation between MFG-1 activation and the proportion of intentional flakes produced during the third session of the Acheulian task indicates that the most skilled toolmakers recruited dlPFC to a greater extent than less skilled toolmakers by the end of the study. Together, these results suggest that WM, and perhaps other executive functions linked to dlPFC, such as planning, reasoning, and inhibition, play an important role in the process of learning to make complex stone tools, specifically early Acheulian handaxes. Although we demonstrate spatial overlap between the current study’s results and known visual WM centres, future replication studies should include WM and other executive function tasks in addition to toolmaking tasks to confirm the colocation of the responses of these tasks within individuals. This would help pinpoint which cognitive functions are involved during these toolmaking tasks.

It is possible that this positive result can be attributed to our use of modern human participants who could possess a derived WM adaptation to solve novel problems (Coolidge & Wynn, 2005). If true, our study would have few, if any, implications for extinct hominin species. However, if this were the case, then there should be similar levels of prefrontal activation during the Oldowan task as well. The fact that this is not observed indicates that dlPFC involvement is contingent upon task complexity, as Acheulian toolmaking is the more complicated of the two tasks.

Both toolmaking tasks are relatively difficult, requiring multiple hours of training to master. For example, our participants completed seven hours of training on these tasks, and none of these participants could be considered expert toolmakers by the end of the experiment. This is reflected in the overall reduction in left dorsal PrG activity over the course of the experiment. The dorsal PrG plays a key role in visually guided reaching, but recent evidence also indicates its involvement during the performance of sequential movements internally generated from memory after extended practice (Ohbayashi et al., 2016). The decrease in activity in dorsal PrG across sessions may indicate that participants began to anticipate upcoming actions as they gained more experience. While this behaviour relies on visual guidance when first being learned, the fact that a higher proportion of flake mass is associated with lower activity in this area after more than seven hours of training suggests that sequential flaking becomes automatized over time. Alternatively, it is possible that activation differences across technologies and between sessions are driven by overt behavioural differences, for example, the number of strikes of the hammer stone against the core or the frequency of amplitude of sounds generated by the knapper’s actions. The total number of debitage elements, a rough proxy for the frequency of striking the core with the hammer stone, does not appear to contribute to neural activation differences. This suggests that the changes in neural activity that we observed can be attributed to internal rather than external factors; however, an event-related analysis based on observable behaviours would shed more light on this issue.

Archaeologists have long suspected that the complexity of Acheulian tools and the procedure involved in their production necessitate a greater degree of cognitive capacity than that required for the Oldowan industry (Stout et al., 2014; Toth & Schick, 2018; Wynn, 1985, 1993). Neuroimaging studies largely support this claim (Putt et al., 2017; Stout & Chaminade, 2007; Stout et al., 2008, 2011, 2015), and the current study demonstrates that this difference in technological complexity is reflected in the neural networks that are involved in learning Oldowan and early Acheulian stone toolmaking skills. Specifically, it appears that Oldowan toolmaking quickly transitions from controlled processing guided by the dorsal visual attention network during the first session to automatic processing in fewer than four hours of training.

The Oldowan task recruits a visual attention network during the first neuroimaging session, involving the inferior parietal lobe and dorsal premotor cortex, regions that are interconnected via recurrent fibres that pass through the superior longitudinal fasciculus (Ptak, 2012). Activated clusters in this network include left MFG-2 and right dorsal PrG. The right MTG and left STG also are recruited. By the second session, both frontal regions and the right MTG become suppressed. In their place arises the right PoG, a primary sensorimotor area, which reaches its peak activation during the second session before decreasing below baseline by the third session. Individuals with the lowest levels of PoG activation in the third session tended to be the most efficient at removing flake mass from a core, which is the main goal of the task. By this point, participants’ Oldowan products resembled those from the archaeological record (Fig. 2a-b), suggesting that the simple removal of flakes without the added element of shaping a core tool was well rehearsed and therefore did not demand active attention. This pattern of deactivation of cognitive control areas and activation of a sensorimotor area is likely the result of a transition to procedural memory after fewer than four hours of practice. Note that deeper brain structures associated with procedural memory, such as the cerebellum (Molinari et al., 1997),

ACCEPTED MANUSCRIPT

could not be recorded using fNIRS to confirm this assertion; however, Stout and Chaminade (2007) report cerebellum involvement during Oldowan toolmaking after four hours of practice.

We did not find a similar pattern of deactivation of cognitive areas and activation of sensorimotor areas over time as participants learned to make early Acheulian handaxes. Rather, the left MFG-2, right PrG (FEF), and bilateral temporal areas increased with additional practice, and activation in the right PoG decreased with more training. This pattern of increasing activation of the control network and coinciding deactivation of a primary sensorimotor area during Acheulian tool production indicates an emphasis on the employment of cognitive strategies at all measured stages of learning. Individuals with the highest levels of PrG (FEF) activation in the third session tended to make the fewest mistakes, in the form of unintentional shatter. Combined with the aforementioned dlPFC activation, these results indicate that handaxe production is likely a visuospatial WM task that consistently engages WM areas, even after multiple hours of training. These results depict early *Homo* as curious, attentive, and capable of some degree of flexible thinking as they learned toolmaking skills.

Toolmaking can be a physically strenuous task, which may influence systemic blood pressure and respiration. Although this study utilized a depth-resolved fNIRS technique (in the form of short channels) to control for extracerebral hemodynamics and tPCA to target and eliminate motion artefacts, systemic signals were not directly measured. Systemic confounds may sometimes lead to false positives and false negatives in fNIRS data (Tachtsidis & Scholkmann, 2016); therefore, this is a limitation of the study. An analysis of HbR in addition to HbO (see Supplementary Materials), however, reveals the typical inverse relationship between the two chromophores. This likely means that systemic signals did not confound the hemodynamic response; nevertheless, the only way to be certain would be to measure systemic signals and neural activity simultaneously.

4.1. Conclusions

We argue that the results of the current study and previous neuro-archaeological experiments (Putt et al., 2017; Stout et al., 2015) support a novel hypothesis: positive selective forces acted on hominin WM networks of the brain as early as 1.8 Ma, when the more complex early Acheulian industry began to emerge in the archaeological record. Under this hypothesis, individuals with derived WM capabilities were the most successful at learning the crucial skills associated with handaxe production. In turn, they and their offspring were more reproductively successful than their counterparts because of the facilitated access to calorically dense and diverse food resources that these tools imparted. It may not be coincidence that a step increase in brain size also occurred around this time (Shultz et al., 2012) that was driven by a disproportionate expansion of the prefrontal and temporal cortices (Bruner & Holloway, 2010), the same areas that are selectively activated by the Acheulian task in the current study. Tool use and WM strongly correlate with brain size (Posthuma et al., 2003; Reader & Laland, 2002). Thus, selection for enhanced WM may have led to an increase in brain size, particularly in the prefrontal and temporal cortices, that occurred near the beginning of the Pleistocene, which set *Homo* on the path to becoming human.

Acknowledgments

ACCEPTED MANUSCRIPT

We thank A. Woods for his time and knapping expertise, J. Pope Johnson for her illustrations in Fig. 2, S. Hayes and S. Forbes for their statistical consulting, and M. Adams, S. Allchin, G. Brua, C. Daniel, E. DeForest, E. Dellopoulos, N. Fox, E. Hoeper, M. Jedele, D. Jones, C. Mundy, A. Vega, and A. Wells for their assistance in the lab. The data collection portion of this project was supported by the Leakey Foundation, the Wenner-Gren Foundation (Grant No. 8968), and Sigma Xi, the Scientific Research Society, while the writing portion of the project was supported by the John Templeton Foundation (Grant No. 52935) and AAUW.

References

Andrefsky, W. *Lithics: Macroscopic Approaches to Analysis* (Cambridge University Press, ed. 2, 2005).

Bamforth, D.B., Finlay, N. Introduction: Archaeological approaches to lithic production skill and craft learning. *J. Archaeol. Method Th.* **15,** 1-27 (2008).

Barbey, A.K., Koenigs, M., Grafman, J. Dorsolateral prefrontal contributions to human working memory. *Cortex.* **49,** 1195-1205 (2013).

Barth, J., Call, J. Tracking the displacement of objects: A series of tasks with great apes and young children. *J. Exp. Psychol.* **32,** 239-252 (2006).

Beyene, Y., *et al.* The characteristics and chronology of the earliest Acheulean at Konso, Ethiopia. *PNAS.* **110**, 1584-1591 (2013).

Bruner, E., Holloway, R.L. A bivariate approach to the widening of the frontal lobes in the genus *Homo. J. Hum. Evol.* **58,** 138-146 (2010).

Calvetti, D., Morigi, S., Reichel, L., Sgallari, F. Tikhonov regularization and the L-curve for large discrete ill-posed problems. *J. Comput. Appl. Math.* **123,** 423-446 (2000).

Carel, C., *et al.* Neural substrate for the effects of passive training on sensorimotor cortical representation: A study with functional magnetic resonance imaging in healthy subjects. *J. Cerebr. Blood F. Met.* **20,** 478-484 (2000).

Carruthers, P. Evolution of working memory. *P. Natl. Acad. Sci. USA.* **110,** 10371-10378 (2013).

Chen, G., Adleman, N. E., Saad, Z. S., Leibenluft, E., Cox, R. W. Applications of multivariate modeling to neuroimaging group analysis: A comprehensive alternative to univariate general linear model. *NeuroImage.* **99,** 571-588 (2014).

Coolidge, F.L., Wynn, T. An introduction to cognitive archaeology. *Curr. Dir. Psychol. Sci.* **25**, 386-392 (2016).

Coolidge, F.L., Wynn, T. Executive functions of the frontal lobes and the evolutionary ascendancy of *Homo sapiens. Camb. Archaeol. J.* **11,** 255-260 (2001).

Coolidge, F.L., Wynn, T. Working memory, its executive functions and the emergence of modern thinking. *Camb. Archaeol. J.* **15,** 5-26 (2005).

Diamond, A. Executive functions. *Annu. Rev. Psychol.* **64,** 135-168 (2013).

Diez-Martín, F., *et al.* Early Acheulean technology at Es2-Lepolosi (ancient MHS-Bayasi) in Peninj (Lake Natron, Tanzania). *Quat. Int.* **322-323**, 209-236 (2014).

Eggebrecht, A.T., *et al*. Mapping distributed brain function and networks with diffuse optical tomography. *Nat. Photonics.* **8**, 448-454 (2014).

ACCEPTED MANUSCRIPT

Fang, Q., Boas, D. Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units. *Opt. Express.* 17, 20178-20190 (2009).

Fritz, C.O., Morris, P.E., Richler, J.J. Effect size estimates: Current use, calculations, and interpretation. *J. Exp. Psychol.* **141**, 2-18 (2012).

Fuster, J.M. Executive frontal functions. *Exp. Brain Res.* **133,** 66-70 (2000).

Gagnon, L., Perdue, K., Greve, D.N., Goldenholz, D., Kaskhedikar, G., Boas, D.A. Improved recovery of the hemodynamic response in Diffuse Optical Imaging using short optode separations and state-space modeling. *NeuroImage*. **56**, 1362-1371 (2011).

Garavan, H., Kelley, D., Rosen, A., Rao, S.M., Stein, E.A. Practice-related functional activation changes in a working memory task. *Microsc. Res. Techniq.* **51,** 54-63 (2000).

Goldman-Rakic, P.S. Cellular basis of working memory. *Neuron.* **14,** 477-485 (1995).

Hardwick, R.M., Rottschy, C., Miall, R.C., Eickhoff, S.B. A quantitative meta-analysis and review of motor learning in the human brain. *NeuroImage.* **67,** 283-297 (2013).

Henshilwood, C.S., Dubreuil, B. The Still Bay and Howiesons Poort, 77-55 ka: Symbolic material culture and the evolution of the mind during the African Middle Stone Age. *Curr. Anthropol*. **41**, 576-590 (2011).

Jansma, J.M., Ramsey, N.F., Slagter, H.A., Kahn, R.S. Functional anatomical correlates of controlled and automatic processing. *J. Cog. Neurosci.* **13,** 730-743 (2001).

Kelly, A.M.C., Garavan, H. Human functional neuroimaging of brain changes associated with practice. *Cereb. Cortex.* **15,** 1089-1102 (2005).

Landau, S.M., Schumacher, E.H., Garavan, H., Druzgal, T.J., D'Esposito, M. A functional MRI study of the influence of practice on component processes of working memory. *NeuroImage.* **22,** 211-221 (2004).

León-Carrión, J., *et al.* Efficient learning produces spontaneous neural repetition suppression in prefrontal cortex. *Behav. Brain Research.* **208**, 502-508 (2010).

Loubinoux, I., *et al.* Within-session and between-session reproducibility of cerebral sensorimotor activation: A test-retest effect evidenced with functional magnetic resonance imaging. *J. Cerebr. Blood F. Met.* **21,** 592-607 (2001).

Molinari, M., *et al.* Cerebellum and procedural learning: Evidence from focal cerebellar lesions. *Brain.* **120,** 1753-1762 (1997).

Moore, C.D., Cohen, M.X., Ranganath, C. Neural mechanisms of expert skills in visual working memory. *J. Neurosci.* **26,** 11187-11196 (2006).

Ohbayashi, M., Picard, N., Strick, P.L. Inactivation of the dorsal premotor area disrupts internally generated, but not visually guided, sequential movements. *J. Neurosci.* **36,** 1971-1976 (2016).

Oldfield, R.C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychol.* **9,** 97-113 (1971).

Perlman, S.B., Huppert, T.J., Luna, B. Functional near-infrared spectroscopy evidence for development of prefrontal engagement in working memory in early through middle childhood. *Cereb. Cortex.* **26**, 2790-2799 (2016).

Petrides, M. The role of the mid-dorsolateral prefrontal cortex in working memory. *Exp. Brain Res.* **133,** 44-54 (2000).

Posthuma, D., *et al.* Genetic correlations between brain volumes and the WAIS-III dimensions of verbal comprehension, working memory, perceptual organization, and processing speed. *Twin Res. Hum. Genet.* **6,** 131-139 (2003).

Ptak, R. The frontoparietal attention network of the human brain: Action, saliency, and a priority map of the environment. *Neuroscientist.* **18,** 502-515 (2012).

Putt, S. S., Woods, A. D., Franciscus, R. G. The role of verbal interaction during experimental bifacial stone tool manufacture. *Lithic Technol.* **39,** 96-112 (2014).

Putt, S.S., Wijeakumar, S. Tracing the evolutionary trajectory of verbal working memory with neuro-archaeology. *Interact. Studies.* **19,** 272-288 (2018).

Putt, S.S., Wijeakumar, S., Franciscus, R.G., Spencer, J.P. The functional brain networks that underlie Early Stone Age tool manufacture. *Nat. Hum. Behav.* **1(0102),** 1-8 (2017).

Read, D.W. Working memory: A cognitive limit to non-human primate recursive thinking prior to hominid evolution. *Evol. Psychol.* **6**, 676-714 (2008).

Reader, S.M., Laland, K.N. Social intelligence, innovation, and enhanced brain size in primates. *P. Natl. Acad. Sci. USA.* **99,** 4436-4441 (2002).

Semaw, S., et al. 2.5-million-year-old stone tools from Gona, Ethiopia. *Nat.* **385**, 333-336 (1997).

Sherwood, C.C., Subiaul, F., Zawadzki, T.W. A natural history of the human mind: Tracing evolutionary changes in brain and cognition. *J. Anat.* **212,** 426-454 (2008).

Shultz, S., Nelson, E., Dunbar, R.I.M. Hominin cognitive evolution: Identifying patterns and processes in the fossil and archaeological record. *Philos. T. R. Soc. B.* **367,** 2130-2140 (2012).

Stout, D. Stone toolmaking and the evolution of human culture and cognition. *Phil. T. Roy. Soc. B.* **366,** 1050-1059 (2011).

Stout, D., Apel, J., Commander, J., Roberts, M. Late Acheulean technology and cognition and Boxgrove, UK. *J. Archaeol. Sci.* **41**, 576-590 (2014).

Stout, D., Chaminade, T. The evolutionary neuroscience of tool making. *Neuropsychol.* **45,** 1091-1100 (2007).

Stout, D., Hecht, E., Khreisheh, N., Bradley, B., Chaminade, T. Cognitive demands of Lower Paleolithic toolmaking. *PLoS ONE.* **10,** e0121804 (2015).

Stout, D., Passingham, R., Frith, C., Apel, J., Chaminade, T. Technology, expertise and social cognition in human evolution. *Eur. J. Neurosci.* **33**, 1328-1338 (2011).

Stout, D., Semaw, S., Rogers, M.J., Cauche, D. Technological variation in the earliest Oldowan from Gona, Afar, Ethiopia. *J. Hum. Evol.* **58**, 474-491 (2010).

Stout, D., Toth, N., Schick, K.D., Chaminade, T. Neural correlates of Early Stone Age tool-making: Technology, language and cognition in human evolution. *Philos. T. R. Soc. B.* **363,** 1939-1949 (2008).

Tachtsidis, I., Scholkmann, F. False positives and false negatives in functional near-infrared spectroscopy: Issues, challenges, and the way forward. *Neurophotonics.* **3**, 031405 (2016).

Tikhonov, A. Solution of incorrectly formulated problems and the regularization method. *Sov. Mathematics-Doklady*. **5,** 1035-1038 (1963).

Toth, N., Schick, K. An overview of the cognitive implications of the Oldowan Industrial Complex. *Azania.* **53,** 3-39 (2018).

Toth, N., Schick, K., Semaw, S. A comparative study of the stone tool-making skills of *Pan*, *Australopithecus*, and *Homo sapiens*. In *The Oldowan: Case Studies into the Earliest Stone Age* (Stone Age Institute Press, 2006).

Van Essen, C., Dierker, D.L. Surface-based and probabilistic atlases of primate cerebral cortex. *Neuron.* **56,** 209-225 (2007).

van Raalten, T.R., Ramsey, N.F., Duyn, J., Jansma, J.M. Practice induces function-specific changes in brain activity. *PloS ONE.* **3,** e3270 (2008).

Vernet, M., Quentin, R., Chanes, L., Mitsumasu, A., Valero-Cabré, A. Frontal eye field, where art thou? Anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations. *Front. Integr. Neurosci.* **8,** 66 (2014).

Washburn, D.A., Gulledge, J.P., James, F., Rumbaugh, D.M. A species difference in visuospatial working memory: Does language link “what” with “where”? *Int. J. Comp. Psychol.* **20,** 55-64 (2007).

Wijeakumar, S., Huppert, T., Magnotta, V. A., Buss, A. T., Spencer, J. P. Validating an image-based fNIRS approach with fMRI and a working memory task. *NeuroImage.* **147,** 204-218 (2017).

Wijeakumar, S., Spencer, J. P., Bohache, K., Boas, D. A., Magnotta, V. A. Validating a new methodology for optical probe design and image registration in fNIRS studies. *NeuroImage.* **106,** 86-100 (2015).

Wynn, T. Piaget, stone tools and the evolution of human intelligence. *World Archaeol.* **17,** 32-43 (1985).

Wynn, T. Two developments in the mind of early *Homo. J. Anthropol. Archaeol.* **12,** 299-322 (1993).

Wynn, T., Coolidge, F.L. Archeological insights into hominin cognitive evolution. *Evol. Anthropol.* **25**, 200-213 (2016).

Wynn, T., Coolidge, F.L. The expert Neandertal mind. *J. Hum. Evol.* **46**, 467-487 (2004).

Yankosec, K.E., Howell, D. A narrative review of dexterity assessments. *J. Hand Ther.* **22,** 258-270 (2009).

Yücel, M.A., Selb, J., Cooper, R.J., Boas, D.A. Targeted principle component analysis: A new motion artifact correction approach for near-infrared spectroscopy. *J. Innovat. Opt. Health Sci.* **7,** 1350066 (2014).

Additional Materials

Author Contributions

SSJP and JPS conceived of and designed the experiment. SSJP and SW collected and analysed the data. SW contributed analysis tools. SSJP drafted the article, with contributions from JPS and SW. All three authors critically revised and approved the final version to be published.

ACCEPTED MANUSCRIPT