METHODS

This chapter focuses on the methods and techniques applied within this project. It is organised according to the different steps of the research and types of analysis. The project is trying to follow an integrative and multidisciplinary approach to gain a more holistic view on asymmetric tools from the Late Middle Palaeolithic. At the same time, a standardised analytical workflow is aimed for. Irrespective of the applied method or pieces of equipment used, following standards and protocols were sought. Used settings, programs and software were reported as transparently as possible. The scripts and templates used to analyse the data are saved and stored on GitHub, a distributed version-control system for tracking changes [https://github.com/lschunk]. The same applies to all involved data, including the raw data.
The first subchapter describes the raw material characterisation, followed by the techno-typological analysis, and a subchapter on the quantification of the edge design. The next two subchapters addressing the use-wear analysis are separated into a qualitative and a quantitative analysis. The last part of this chapter is dedicated to the controlled experiments. All methods are targeted at gaining information about the archaeological assemblages and producing data, which in most cases is statistically evaluable. Although it goes beyond the scope of a method chapter, an introduction to some of the methods are given. This is the case for the chapters addressing raw material characterisation, use-wear analysis, edge design and experiments. Thereby, the introductions focus on the perspective of lithic analysis. Additionally, the advantages, disadvantages and limitations of these methods are discussed.

RAW MATERIAL CHARACTERISATION

Early humans had to select suitable raw material for the production of tools. Whether this process was a conscious decision or not was also presumably dependent on the raw material availability and variability. Research on the raw material variability present in archaeological assemblages has already been established for many decades (Dibble 1985; Féblot-Augustins 1993; Floss 1994; Féblot-Augustins 1997; 2008; Andrefsky 2009; Meignen/Delagnes/Bourguignon 2009). The reasons for studying the raw material can differ. The most obvious one is the classification of the raw material and identification of the possible source, which can provide information about movement patterns of early hominids (Andrefsky 1994; Féblot-Augustins 1997; Brantingham 2003; Féblot-Augustins 2008). Alongside raw material procurement, the raw material properties themselves are another aspect which can be analysed. The properties of raw material, for instance the size and shape of the original blocks or nodules as well as density and hardness, have a critical influence on the ability to knap a tool (Kuhn 1992; Braun et al. 2009; Dogandžić et al. 2020). Moreover, the tool design correlated with the used raw material can provide information about intentional selection. Therefore, raw material properties are assumed to affect not only tool production, but also technical choices in general and tool use specifically (Odell 1981; Delgado-Raack/Gómez-Gras/Risch 2009; Nonaka/Bril/Rein 2010; Key/Proffitt/de la Torre 2020). By the analysis of their characteristics, it is possible to say whether, for example, physical properties have been manipulated through heat treatment (Brown et al. 2009; Schmidt/Mackay 2016; Key/Pargeter/Schmidt 2020). Features such as surface roughness are also recognised as a determining factor in the development and the characteristics of use-wear traces (Lerner 2014). As demonstrated by several studies, the nature and the properties of the contact material can be associated with specific

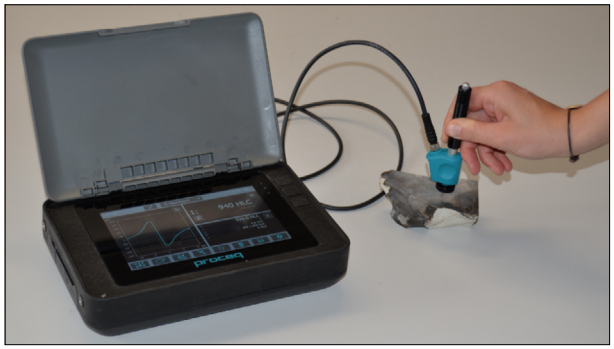


Fig. 19 Hardness measurements taken on a flint sample using the Leeb rebound hardness tester with the Leeb C probe (Proceq Equotip 550).

use-wear traces (Stemp/Stemp 2003; Lerner 2007; 2014). The rate at which use-wear forms on a given tool surface, is dependent on the surface hardness and roughness of both the tool and the contact material. This means that raw material properties are known to affect the development and the characteristics of use-wear traces, for instance polish (Lerner 2007; 2014; Lerner et al. 2014). Analysing the material properties of a given raw material is therefore a prerequisite to understand the formation of use-wear traces and their interpretation, respectively. To go into more detail, raw material properties are an influencing factor not only for tool morphology, but also for tool efficiency and durability. As demonstrated recently, early hominins might have actively chosen certain raw materials for specific functional performances (Agam/Zupancich 2020; Key/Proffitt/de la Torre 2020). Consequently, analysing raw material properties should be part of each lithic study.

Thus, the first aspect of this study addresses the raw material used to produce the tools. In the case of the three examined sites, only two types of raw material are involved in the analysis: silicified schist and cretaceous (»Baltic«) flint (Vandebosch 1921; Günther 1964; Bosinski 1969; Jöris 2001). For the petrographic analysis, only optical factors such as the colour, the nature of the cortex and the translucency of the raw material were of relevance. The raw material from the archaeological sites was compared with the lithic raw material collection at MONREPOS Archaeological Research Centre and Museum for Human Behavioural Evolution, RGZM as a reference (https://monrepos.rgzm.de/en/ausstattung/#Sml). This collection contains siliceous rock samples from different sources used to manufacture tools in western Central Europe during the Palaeolithic and Mesolithic (Floss 1994).

The properties of the raw materials, for instance hardness, can be measured to evaluate the possible influence of the raw material itself on tool morphology and also to predict potential efficiency and durability during the use-life of the tool (Key/Lycett 2017; Key/Pargeter/Schmidt 2020; Key/Proffitt/de la Torre 2020). Relatively common is the use of a Schmidt hammer to estimate mechanical properties of rock material (Yilmaz/Sendir 2002; Yaşar/Erdoğan 2004; Braun et al. 2009). The Schmidt hammer is a device that works with a rebound hammer to test surface hardness and penetration resistance based on elastic properties and the strength of rock. An alternative method has been provided recently using a Leeb rebound hardness tester. The hardness test was originally developed for metals but has been correlated for rocks (Braun et al. 2009; Rodríguez-Rellán 2016; Corkum et al. 2018; Kovler/Wang/Muravin 2018). The measurements can be carried out rapidly and non-destructively. Micro surface roughness of knapped surfaces as another parameter, can be measured for example with a 3D confocal laser scanning microscope. Even though important, the material properties of the raw materials present in archaeological assemblages have been rarely tested. Although highly relevant in experimental design, they are often neglected (Evans et al. 2014). The reason for this might be the costly equipment needed to calculate some of the properties or maybe an underestimation of their relevance.

Here, the material properties of the raw material were tested using the Leeb rebound hardness tester (Proceq Equotip 550, Leeb C probe; **fig. 19**). Two types of raw material were involved in the analysis: silicified schist and Baltic flint. The advantage of the hardness test conducted with the Leeb rebound is that the measurements can be carried out rapidly and are non-destructive. The disadvantages are the required flat surface and mandatory maximum size of the object. For the knapped archaeological tools, it was not possible to measure the hardness with the Leeb rebound. Based on their morphology it is impossible to position

the tools entirely flat, so that one face is in contact over its total surface with the stable supporting base (here a flat rock slab of about 20 kg). Additionally, the other tool surface must be flat as well in order to conduct the measurements. However, one semi-finished tool from Balve (ID MU-278) fulfilled these requirements. While this approach was used only once for the archaeological samples, it was always applied for the experimental samples (see chapter Controlled experiments). With the hardness tester, to insure and test intra-variability in each sample, ten measurements per sample were taken. The number of measurements not only ensures the inclusion of the raw material variability within the result but also identifies potential outliers. When the samples did not fulfil minimum sample size requirements, an additional coupling paste was used to connect the sample with a massive support plate.

TECHNO-TYPOLOGICAL ANALYSIS

Initial information for all assemblages from the three different archaeological sites was acquired through traditional lithic analysis methods (Hahn 1993; Debénath/Dibble 1994; Andrefsky 1998; Odell/Vergès 2014). This technological and typological analysis was based on two different but complementary methods. While one approach produced qualitative data related to several descriptive attributes, e. g. tool morphology, the shape and character of the back or the application of the *Prądnik method*, the second method generated quantitative results obtained by standard calliper-based morphometric data (e. g. length, width and thickness). *Keilmesser* as well as *Prądnik scrapers* were analysed in the exact same way. The approach described here is therefore identical for both artefact categories respectively. Mentionable exceptions are the *Prądnik spalls*, the *Keilmesser tips*, scrapers and simple flakes, which are treated differently in some respects. These artefact categories were analysed in the same way but certain steps of the analysis (e. g. definition *Keilmesser* shapes, perimeter sections) were not of relevance and have therefore been omitted. Concerning the »*Keilmesser tips*«, it needs to be noted that these tools are notably shorter than a complete *Keilmesser*. Thus, their metric values are not directly comparable with those of the remaining artefacts.
The database for the techno-typological analysis was done by using E4 for Microsoft Windows 10 (Old Stone Age, S. McPherron), as a data entry program. To use E4, a script with all variables was entered and their type (e. g. text, numeric) was configured. Since E4 offers the possibility, to create conditional statements, variables can be skipped based on values entered for previous variables. The script was written in the open-source software R (version 4.0.2 through RStudio version 1.3.1073, RStudio Inc., Boston, USA) for Microsoft Windows 10. The data was saved in MDB format.

Blanks and cortex

To begin with, a qualitative attribute is addressed with the determination of the blank used for the tool production. The tools can be manufactured on different blanks (Jöris 2001). If a distinction is possible, it should be determined in this part of the analysis, whether a core or a flake was used as a blank to produce the tool. In some cases, it is not possible to identify the type of blank. Thus, the term »undeterminable« was selected. Another feature that can provide valuable information, for instance on manufacturing sequences or methods, is the amount of cortex left on the tool surface (Dibble et al. 2005). The first question is whether there is cortex on the artefact. If yes, the amount of cortex in percent (>25 %, 25-50 %, 50-75 %, <75 %, total) was indicated as well as the location.

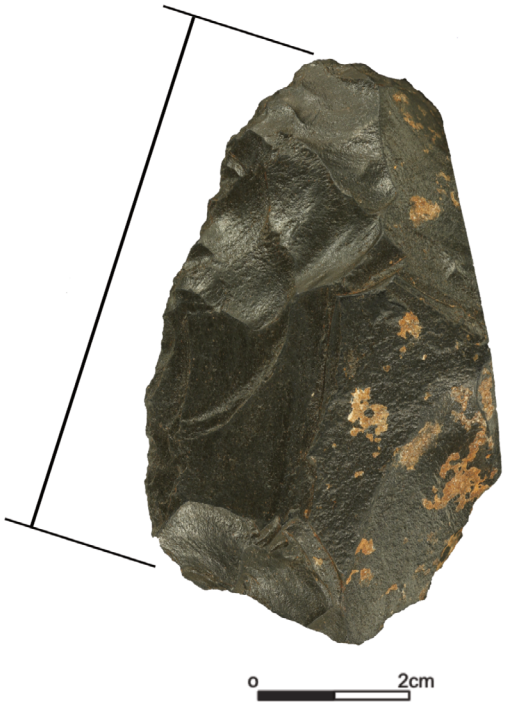


Fig. 20 *Keilmesser* from Balver Höhle (ID MU-021). The black lines indicate the length measurement of the active edge. – Scale 1:1.

Morphometric quantitative analyses

All metric data were recorded using a digital calliper of the brand Mitutoyo with a direct data output. The used model achieves a precision of 0.01 mm with an error margin of 0.03 mm. Dimensions were taken for complete and incomplete artefacts, the latter being recoded as such. The length measurement was determined using the active edge and not necessarily the maximal dimension of the tool (fig. 20). However, the maximum dimensions were applied for the measurements of width and thickness (Debénath/Dibble 1994).

Tool back

The design of the back forms another of the criteria of the attribute analysis. Here, the relevant trait is the state of the back (Bosinski 1967; Jöris 2001). The design of the artefact back can inform about the extent to which the back already formed a conceptual part of the tool manufacture and influence of the morphology of the raw material blank used. For instance, the back can be characterised by unworked cortex surfaces or natural fragmentation. Moreover, clear traces of intentional retouch or minor working traces can be observed.
The thickness of the tool back was recorded identically to the dimensions described above, at its maximum extent, meaning the most pronounced section of the back.

Active edge retouch

In addition to the back, the design of the active edge was analysed. The interesting technological aspect was the type of retouch along the edge. The retouch was classified into either bifacial retouch, semi-bifacial retouch or unifacial retouch.

Lateralisation

Another feature documented as a qualitative attribute is the laterality of the artefacts. The lateralisation results from tool asymmetry (Jöris/Uomini 2019; fig. 9). Tool lateralisation can be determined when the artefact is aligned with the proximal part of the tool towards the observer and the flat surface (of *Keilmesser*) or the ventral surface (of *Prądnik scrapers*). If the back is then on the left side and the active edge on the right side, the artefact counts as *dextrolateral* (dex.). If the back is to the right and the active edge at the left, the tool is recorded as *sinistrolateral* (sin.).

Morphological shape variability

The artefacts were assigned to one of the seven distinct *Keilmesser* shapes respectively (**fig. 7**). However, the division was made baes only on a visual impression. The fundamental question to be asked is whether these *Keilmesser* shapes indeed exist as strictly distinguishable types or if the types merge into each other and show the result of reworking and resharpening processes. In the following, the seven *Keilmesser* shapes are described in more detail (e. g. Jöris 2001).

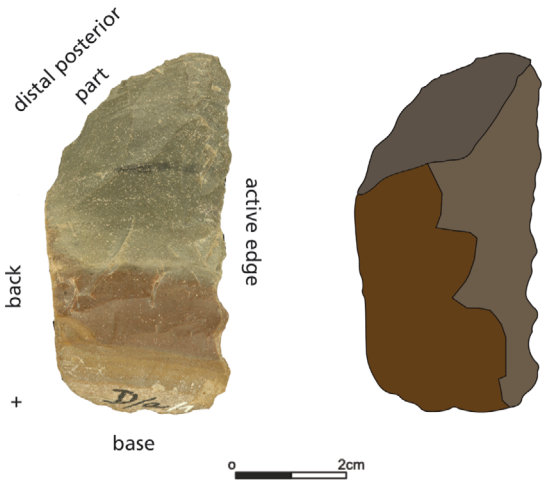


Fig. 21 Dorsal surface of a *Keilmesser* from Balver Höhle (ID MU-114) (left), indicating the different perimeter sections base and back, distal posterior part as well as active edge. The half-schematic illustration (right) displays the identical *Keilmesser*. The three perimeter sections are highlighted in colours. – Not to scale.

The »*Bockstein knife*« is characterized by the absence of a transition between the back and the active edge. Both these tool areas, the back and the active edge, meet at an acute-angle (Wetzel/Bosinski 1969).

The back and the active edge of the »*Klausennische Messer*« run almost parallel to each other. A bifacially worked, mostly obliquely running distal posterior part connects the back with the active edge. This distal posterior part and the active edge display an acute angle (Wetzel/Bosinski 1969).

The »*Prądnik Messer*« is morphologically closest to given definition of a »*Klausennischen Messer*«. The back of this *Keilmesser* shape changes into a retouched, bow-shaped distal posterior part, which is more rounded and connects with the active edge at approximately right angles (Wetzel/Bosinski 1969).

The outline of the »*Buhlener Keilmesser*« resembles a right-angled, equilateral triangle. The two catheti are formed by the back and the distal posterior part, whereas the active edge corresponds to the hypotenuse (Jöris 2001).

Characteristic for the »*Lichtenberger Keilmesser*« is a clear convexity of the active edge. Both the base and the distal posterior part are rounded and thus the active edge is slightly elongated. The description of this *Keilmesser* shape is not narrowly defined and can vary (Veil et al. 1994).

Artefacts of the type »*Königsaue-type Keilmesser*« are also characterized by a convex active edge. The back is located at the proximal part of the tool. The biconvex distal end of the artefact resembles the morphology of a leaf point (Mania/Toepfer 1973).

The »*Balver Keilmesser*« is defined by a crudely worked base. From the top view, the base seems rounded. The distal posterior part runs obliquely, comparable to the »*Klausennische*« *Keilmesser* shape (Jöris 2001).

Following this descriptive morphology of *Keilmesser*, the outline of the artefacts can be separated into three distinct components, resulting together in the artefact's perimeter (Jöris 2001; Frick/Herkert 2019; **fig. 21**). Firstly, one segment is defined by the tool base and the back. The second section is described by the distal posterior part, connected directly to the back. The overall perimeter is completed by the third region made up of the active edge of the tool. The respective lengths of the different sections of the artefact's perimeter were measured. Only complete tools were suitable for this measurement.

The morphological distinction, of the *Keilmesser* shapes defined above is largely a reflection of the different length ratios of the individual perimeter sections (Jöris 2001; Pastoors 2001; Jöris 2004; Migal/Urbanowski 2006). The metrical measurements can therefore as an objective control whether the different subjectively defined *Keilmesser* shapes exist as valid entities or not. If they are confirmed as valid the different *Keilmesser* shapes should cluster as such when the metrical results are plotted.

Application of the *Prądnik method*

Use of the *Prądnik method* leaves a characteristic negative on the surface of the artefact (Jöris 1992; 2001; Frick at al. 2017; Frick/Herkert 2019; Frick 2020a; fig. 4). The purpose of this part of the analysis is to document whether the method has been applied or not. In some cases more recent negatives overlap with negatives resulting from a previous application of the *Prądnik method.* In this case, the information was labelled as »multiple« applications. Artefacts for which the presence of a negative is debatable are classified as »undetermined«.

Prądnik spalls

Lateral *Prądnik spalls* emerge during the artefact modification by one or more blows to the distal part of the active edge. It is possible to distinguish primary from secondary sharpening spalls based on the pattern of scars on the dorsal face of the flake (Jöris 1992; 2001; Frick/Herkert 2019; fig. 5). The lateralisation, as described before, can also be determined for the *Prądnik spalls*. Based on the orientation of the retouch it can be defined if the spall has been removed from a left-sided or right-sided tool (Cornford 1986; Jöris/Uomini 2019).

Data analysis

All descriptive analyses (summary statistics, scatter plots and bar plots) were performed in the open-source software R version 4.0.2 through RStudio version 1.3.1073 (RStudio Inc., Boston, USA) for Microsoft Windows 10. The following packages were used openxlsx v. 4.1.5, R.utils v. 2.9.2, doBy v. 4.6.7, ggsci v. 2.9, dplyr v. 1.0.2, patchwork v. 1.0.1, ggplot2 v. 3.3.2, tidyverse v. 1.3.0, wesanderson v. 0.3.6. The analysis was done for each archaeological site individually and once for the three sites together. Reports of the analysis in HTLM format, created with knitr v. 1.29 and rmarkdown v. 2.3 are available on GitHub [https://github.com/lschunk/Lithic_analysis_archaeology]. Moreover, the raw data, the scripts and the RStudio project are saved in the same repository.

Sampling strategy

During the techno-typological analysis, samples of the three assemblages were selected for further examination. Artefacts were chosen for edge angle measurements as well as for the use-wear analysis.
Use-wear analysis is a time-consuming methodology, in most cases necessitating sub-sampling of the assemblage. Depending on the research question, the sampling can follow different strategies (Marreiros et al. 2020; fig. 22). What often happens though, is that the supposedly more important artefacts, for instance *fossiles directeurs* or retouched tools will form most of the subset. While this sampling approach would be still debatable, the following interpretation is mostly not. The results obtained for this chosen subset are then applied to the entire assemblage. Sampling based on this rather problematic strategy is frequently criticised as biased. To avoid this bias, sampling strategies should be defined by the research question and not be influenced by preconceived interpretations. The intention is to transfer the analytical results for a subsample to the total assemblage. Elements of all represented technological categories should be integrated

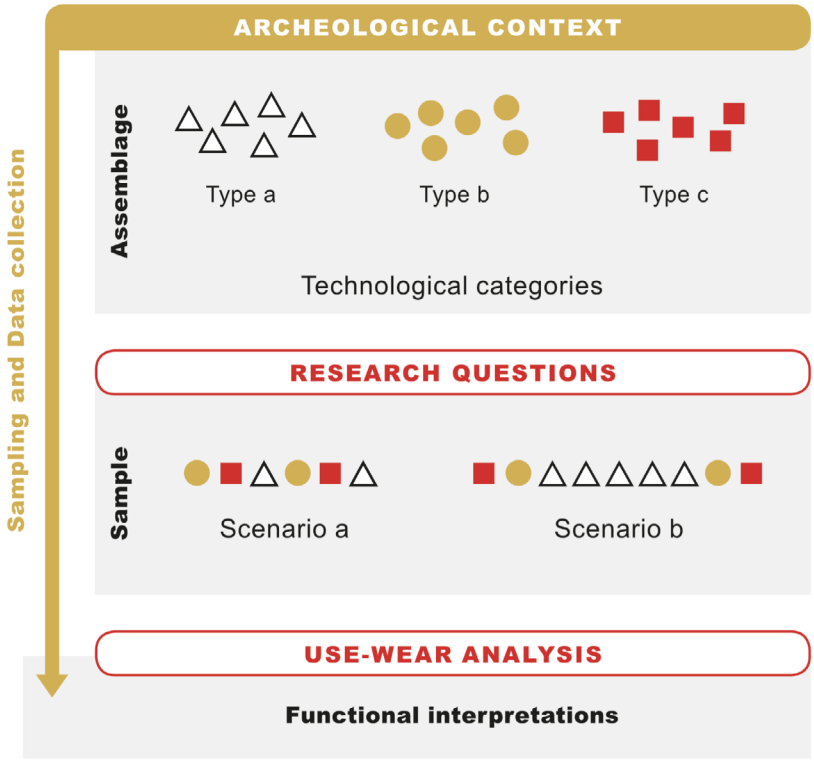


Fig. 22 Schematic representation of the different sampling and data collection strategies commonly used in functional studies. Scenario b was chosen as a sampling strategy for the presented project. – (After Marreiros et al. 2020).

in the analysis. This includes retouched artefacts as well as unretouched specimens often ignored because believed to be unused. If the research question focusses on the use of a certain artefact type, this specific artefact type should represent the most numerous group in the analysis, but outgroups should also be selected to form part of it. Otherwise, the data do not allow for any comparison. Following this idea, lithics were sampled comprehensively to represent all the existing categories present (*Keilmesser*, *Keilmesser tips*, *Prądnik scrapers* and *Prądnik spalls*) as well as *Keilmesser* shapes from the two different raw materials. Additionally, scrapers and flakes were selected as outgroups. Crucial for the sampling was also the preservation of the artefacts.

With this sampling strategy as a prerequisite, artefacts have been selected for two purposes: firstly, for the quantification of their edge design and secondly, for the qualitative use-wear analysis. The sub-sample for calculating the edge angle consists of n = 226 artefacts (**tab. 8**). Only complete artefacts were selected.

site		artefact category						
		Keilmesser	*Keilmesser* tip	*Prądnik scraper*	*Prądnik spall*	scraper	flake	total
Buhlen	n	83 [115]	8 [15]	16 [24]	21 [42]	2 [2]	0 [0]	130 [198]
	%	63.9	6.2	12.3	16.2	1.5	0.0	100.0
Balve	n	65 [170]	10 [21]	3 [27]	0 [117]	0 [12]	0 [0]	78 [347]
	%	83.3	12.8	3.9	0.0	0.0	0.0	100.0
Ramioul	n	9 [9]	0 [0]	1 [3]	0 [0]	6 [6]	2 [2]	18 [20]
	%	50.0	0.0	5.6	0.0	33.3	11.1	100.0
total	n	157 [294]	18 [36]	20 [54]	21 [159]	8 [20]	2 [2]	226 [565]
	%	69.2	7.9	8.8	9.3	3.9	0.90	100.0

Table 8 Number of artefacts from Buhlen, Balver Höhle and Ramioul selected for the quantification of the edge design. The numbers in the square brackets sow the total number of assemblage, respectively.

site		artefact category						
		Keilmesser	*Keilmesser* tip	*Prądnik scraper*	*Prądnik spall*	scraper	flake	total
Buhlen	n	58 [115]	8 [15]	15 [24]	18 [42]	2 [2]	0 [0]	**101** [198]
	%	57.4	7.9	14.9	17.8	2.0	0.0	100.0
Balve	n	42 [170]	2 [21]	5 [27]	21 [117]	9 [12]	0 [0]	**79** [347]
	%	53.2	2.5	6.3	26.6	11.4	0.0	100.0
Ramioul	n	9 [9]	0 [0]	3 [3]	0 [0]	6 [6]	2 [2]	**20** [20]
	%	45.0	0.0	15.0	0.0	30.0	10.0	100.0
total	n	**109** [294]	**10** [36]	**23** [54]	**39** [159]	**17** [20]	**2** [2]	**200** [565]
	%	54.5	5.0	11.5	19.5	8.5	1.00	100.0

Table 9 Number of artefacts from Buhlen, Balver Höhle and Ramioul selected for the qualitative use-wear analysis. The numbers in the square brackets sow the total number of assemblage, respectively.

These artefacts can be divided into n = 157 *Keilmesser*, n = 18 *Keilmesser* tips, n = 20 *Prądnik scraper* and n = 21 *Prądnik spalls*. Additionally, artefacts not belonging to the selected artefact categories were samples for references as a relevant outgroup. Thus, n = 8 scrapers as well as n = 2 flakes.
For the use-wear analysis, more than one third of the total assemblage has been sampled (**tab. 9**). Concretely, n = 200 artefacts from the three archaeological sites have been selected. With n = 129 artefacts, mainly *Keilmesser* (including n = 10 *Keilmesser* tips) have been sampled. Additionally, n = 23 *Prądnik scraper* as well as n = 39 *Prądnik spalls* have been chosen for the qualitative use-wear analysis. Also here, n = 17 scraper and n = 2 flakes build a typological outgroup within the analysis.

QUANTIFICATION OF EDGE DESIGN

Tool design incorporates several attributes, from the overall morphology through to a specific edge retouch. Thus, design builds the bridge between tool shape on the one hand and function and handling on the other. Several methods exist to analyse different aspects of lithic tool design. Recording metric variables, for instance the tool dimensions and mass, presents the most standard of these (Kuhn 1990; Hahn 1993; Andrefsky 1998; Dibble et al. 2017). Although useful, orthogonal dimensions do not capture the geometric configuration of the artefact since artefacts can in fact have similar linear measurements but possess have different morphologies (Borel/Dobosi/Mocel 2017). More detail can be gained through methods such as a geometric morphometrics (e. g. Lycett/Chauhan 2010; Shott/Trail 2010; Serwatka 2015; Weiss et al. 2018) or elliptical Fourier analysis (e. g. Iovita 2009; 2010; Archer et al. 2018; Wiśniewski et al. 2020). Geometric morphometrics focuses on the shape of the tool. Based on defined landmarks and semi-landmarks as well as the spatial and geometric relation between them, this approach captures morphologically distinct shape variables. Elliptical Fourier analysis follows the same main idea to capture the relation between size and shape. It was designed to describe shape by a closed outline, independent of the object's orientation. The method does not depend on landmarks, but on the elliptical Fourier coefficients. The so-called *techno-functional analysis* (Lepot 1993; Boëda 2001; Boëda/Auduze 2013) tries to combine information about the tool morphology and the location of modified and unmodified tool areas. This approach aims to understand tool design including tool function.

Other analyses go further into detail and focus directly on specific areas of the artefact, for example retouched edges. Important therefore is the recognition of an artefact not as a morphological end product (Dibble 1987; Rolland/Dibble 1990; Dibble 1995; Iovita 2014). Tool use as well as maintaining processes can change a tool's shape. Based on retouch intensity, assumptions about the length of the tool-life history are possible (Lin/Marreiros 2020). Identifying these reduction dynamics (Kuhn 1990; Morales/Lorenzo/Vergès 2015) is important to make behavioural inferences based on tool design. As an integrated part of the overall tool design, the tool edge, whether modified or unmodified, can be analysed by a qualitative description (e. g. »stepped scaler retouch«, »trapezoidal retouch«) or quantified by calculating the edge angle. The performance of a given task with a tool is linked with its edge angle (Collins 2008; Key/Lycett 2015; Key/Lycett 2020; Key/Proffitt/de la Torre 2020). This is based on the assumption that the presence of sharp active edges determines a stone tool's utility, whereas the form of a tool's edge determines its functional efficiency, effectiveness, reliability and durability (Key/Lycett 2014; Key/Proffitt/de la Torre 2020). Therefore, the acuteness of an edge angle often plays a role in the interpretation of a tool's function. Several techniques for the recording of measurements have been devised. Among them are manual techniques (e. g. Dibble/Bernard 1980; Pop 2013) or technically more advanced methods, using 3D models (Zaidner/Grosman 2015; Archer et al. 2018; Weiss et al. 2018; Stemp/Macdonald/Gleason 2019; Porter/Roussel/Soressi 2019; Valletta et al. 2020; Weiss 2020; see also ISO 8442-5), all of which being afflicted with their own levels of imprecision inherent to the applied techniques. The recent developments with the integration of script-, algorithm-based and automated analysis of 3D models are clear improvements. On the one hand, it increases the possibility to gain considerable information, and leads to reproducible results and on the other hand decreases random errors. Nevertheless, such methods often result in an overall numeric value per artefact, ignoring possible variability along the edge. To address and understand tool design, a mean value of the edge angle is not sufficient. With more detail, it is possible to address technical choices as well as the effects of retouch on the tool. Thus far, the recording methods are limited by the resolution and accuracy.

Edge angle variability and morpho-functional design

One integrated part of the overall tool design is the tool's active edge. It is common to describe the active edge in a qualitative way (e. g. sharpness, retouch) or quantified by calculating the edge angle (Dibble/Bernard 1980; Pop 2013; Zaidner/Grosman 2015; Archer et al. 2018; Stemp/Macdonald/Gleason 2019; Porter/Roussel/Soressi 2019; Valletta et al. 2020; Weiss 2020). The edge angle describes the angle between two intersecting planes. In the case of lithics, it means the intersection between the dorsal and the ventral surface. To address and understand tool design, it is a prerequisite to know the edge angle values whereby a mean value is not sufficient. With more detail, it is possible to address technological choices as well as the implications of retouch on the tool (Dibble 1995). To acquire the data, an objective method to measure the edge angle at cross sections along the entire tool edge in defined steps as well as measurements at different distances perpendicular to the edge, was chosen. The method is applicable with a 3D model, script based and semi-automated. It provides a systematically error avoiding and reproducible approach, resulting in quantitative and statistically evaluable data.

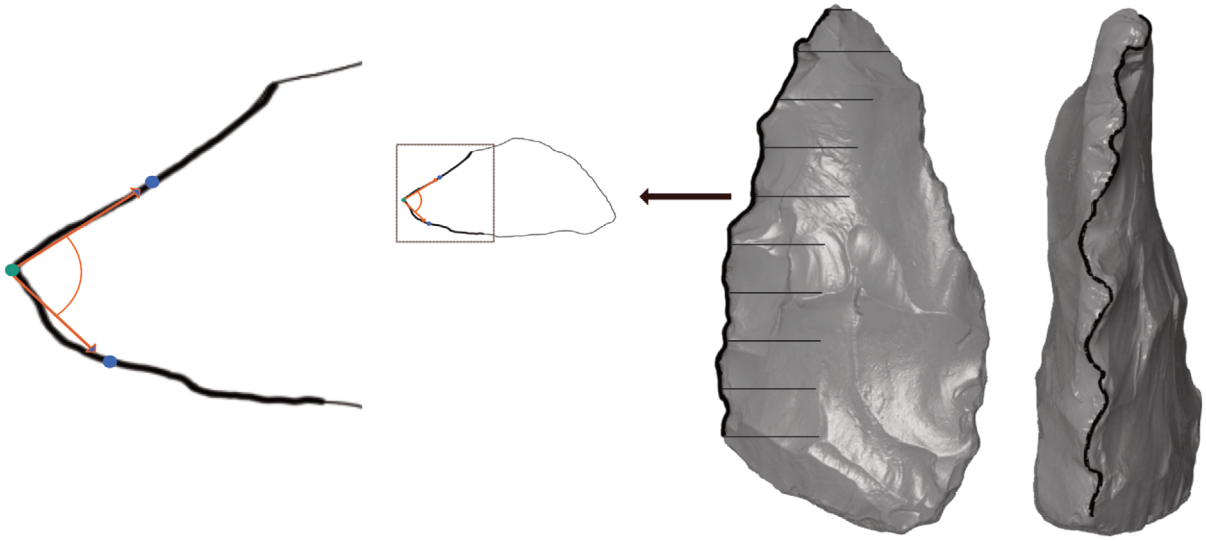


Fig. 23 3D scan of a *Keilmesser* from Ramioul (ID R-007). The scan displays the ventral surface (left) and the active edge (right). The vertical irregular black line represents the »surface curve«, the ten horizontal black lines the sections. The left part of the illustration shows the »3-point« measurement procedure explained on the cross-section of the *Keilmesser* at the position of the fifth section. The green point indicates the intersection between the vertical »surface curve« and the horizontal section. The two blue points are placed on the section in a defined distance. The edge angle measurement (orange lines) is taken based on these three points.

3D data preparation and acquisition

The application of the method is based on the use of a 3D model. Thus, the selected samples were scanned with an AICON smartScan-HE R8 from the manufacturer Hexagon, Germany (software version OptoCat 2018R1). The S-150 FOV used has a resolution of 33µm. Identical settings were used for all scans. Approximately 20 to 27 scans per tool were needed to create a closed model – a prerequisite for the application of the method. Afterwards, the scans were exported in an STL-format.

With a complete model as a basis, some preparatory steps were necessary before calculating the edge angle. The following steps were manually done in GOM Inspect, an open source software for 3D measurement data (GOM Software 2018, Hotfix 1, Rev. 111729, build 2018-08-22). The edge of interest for calculating the edge angle needed to be defined. Since archaeological tools are variable in themselves and display a complex geometry, this is a step that has to be done manually to reach the intended accuracy. Therefore, a »surface curve« was added to each 3D model (**figs 23-25**). This means, a polyline was defined, tracking the edge of the object as precisely as possible and exported in an IGES-format. This step was repeated for all tool edges. In case of *Keilmesser*, most of the time, three polylines (one for the active edge and two for the back) had to be digitalised. With this procedure, the interpretation should be kept apart from the analysis.

The next steps are script-based and also done in GOM Inspect. Based on the defined polyline the software replaces this line with a »reference curve«, which is slightly straightened by filtering the changes in direction out. It is a smoothing function, reducing the point distance to one mm. This parameter was chosen according to the resolution of the 3D model and therefore not a universal parameter. The »reference curve« was used for only one purpose, which is the definition of the sections. The sections are the horizontal lines originating rectangular from the »reference curve« and following the artefact surface in a defined length. These sections are horizontal polylines compared to the vertical polylines along the edge. In case of using the »surface curve« instead of the »reference curve« polyline, the sections would not always be rectangular orientated because of the small changes in direction as a result of the detailed and complex artefact surface.

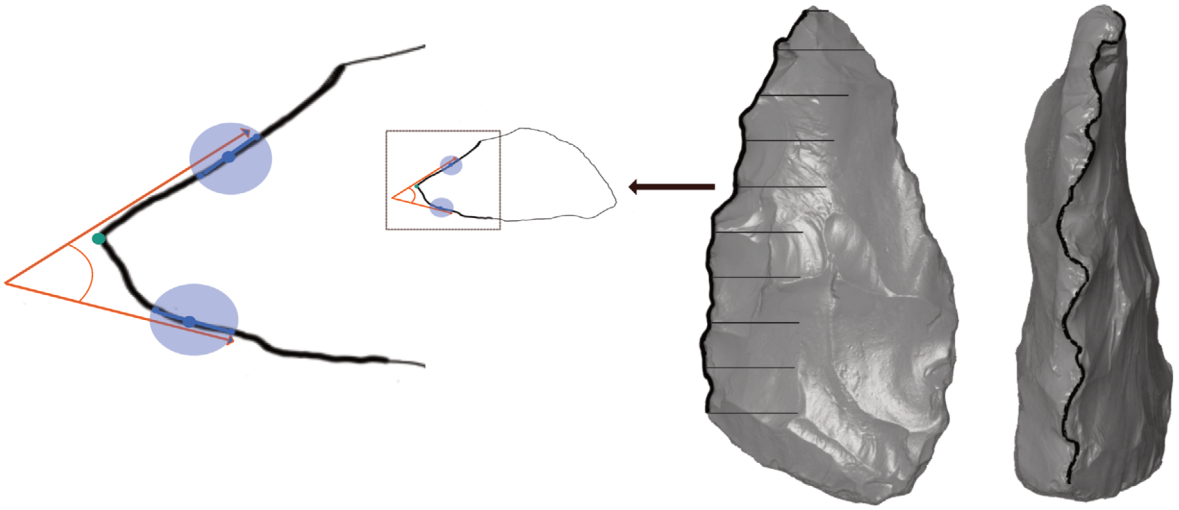


Fig. 24 3D scan of a *Keilmesser* from Ramioul (ID R-007). The scan displays the ventral surface (left) and the active edge (right). The vertical irregular black line represents the »surface curve«, the ten horizontal black lines the sections. The left part of the illustration shows the »2-lines« measurement procedure explained on the cross-section of the *Keilmesser* at the position of the fifth section. The green point indicates the intersection between the vertical »surface curve« and the horizontal section. The two blue points are placed on the section in a defined distance. They build the centre for the two lines. The edge angle measurement (orange lines) is taken based on intersection and the constructed lines.

After generating the »sections«, the »reference curve« was no longer needed. All further steps used the previously created »surface curve«.

»3-point« procedure

For calculating the edge angle in a final step, three different measuring procedures are possible (**fig. 23**). The first one, the so-called 3-point measurement, is modelled on the previously mentioned »*caliper method*« described by H. Dibble and M. Bernard in 1980. Whereas the result of the »*caliper method*« is based on a measurement taken by a special modified calliper at a known distance from the edge of the artefact by using a formula (Dibble/Bernard 1980), the 3-point measurement is using the same topographical indexes represented by 3 points. One of the points is the intersection between the vertical polyline, the »surface curve«, and the horizontal polyline, the section. The two other points are placed on the section, each on one surface, in a defined distance away from the intersection. The edge angle can be calculated between these three points.

»2-lines« procedure

The second measuring procedure is the so-called 2-lines measurement (**fig. 24**). This approach takes the intersection between the two polylines, as described in the first method, as one point. Beginning with this point, it follows the section on both surfaces in a given distance. Until this step, the two procedures are identical. These points on the section are used as a reference for two constructed lines. The lines have a defined length and take the points as a centre from where they spread in both directions. The calculation of the edge angle takes the two lines and the intersection into account.

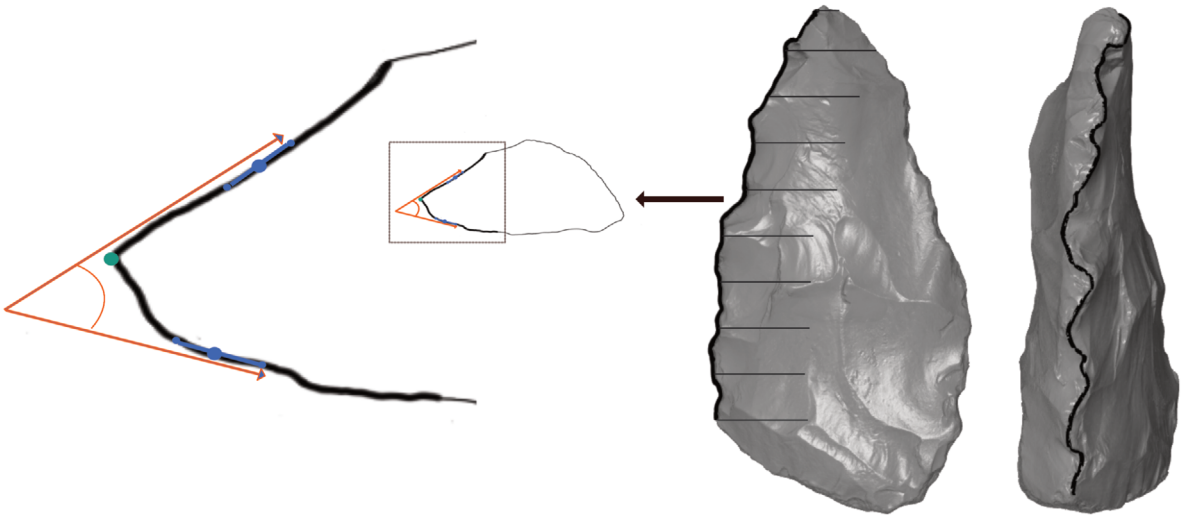


Fig. 25 3D scan of a *Keilmesser* from Ramioul (ID R-007). The scan displays the ventral surface (left) and the active edge (right). The vertical irregular black line represents the »surface curve«, the ten horizontal black lines the sections. The left part of the illustration shows the »best-fit« measurement procedure explained on the cross-section of the *Keilmesser* at the position of the fifth section. The green point indicates the intersection between the vertical »surface curve« and the horizontal section. The two blue points are placed on the section in a defined distance. They build the centre for the two lines. All mesh points of the 3D model in between the two lines are interpolated (indicated by the light-blue circles). The edge angle measurement (orange lines) is taken based on intersection and the constructed lines.

»best-fit« procedure

This approach – the so-called »best-fit« – takes the intersection between the two polylines, the vertical (surface curve) and the horizontal (section) polyline, as one point (**fig. 25**). Beginning with this point, it follows the section on both surfaces in a given distance. These points on the section are used as a reference for two constructed lines. The lines have a defined length and take the points as a centre from where they spread in both directions. The mesh points between the constructed lines were interpolated. The calculation of the edge angle takes the two lines and the intersection into account.
Since it seems like the »best-fit« procedure is more precise and accurate for complex archaeological artefacts, a bit more emphasis was put on these results in the following data analysis.

Parameters

There are several parameters or variables, which need to be chosen in order to apply the method. One of these variables is the number of horizontal sections, represented here by ten sections per tool. The length of the lines, starting at the intersection, was defined with 12 mm. Another parameter is the distance from the edge along the sections. In the described procedures, the intersection between the vertical polyline and the horizontal section determine one crucial point, but the other two (always identical) points per measurement have to be defined numerically. Here the points were defined in 1 mm steps. The last variable is the lengths of the lines used in the »2-lines« and »best fit« measuring procedure, which were always 2 mm.

Edge angle data analysis

The obtained, calculated edge angle values, were analysed. This enabled a comparison between artefacts. To do so, a statistical analysis was performed with R (version 4.0.2 through RStudio version 1.3.1073, RStudio Inc., Boston, USA). The following packages were used: writexl v. 1.3, tidyverse v. 1.3.0, openxlsx v. 4.1.5, ggrepel v. 2.3, doBy v. 4.6.7, patchwork, v. 1.0.1, ggplot2 v. 3.3.2 and R.utils v. 2.9.2. Reports of the analysis in HTLM format, created with knitr v. 1.29 and rmarkdown v. 2.3 are available on GitHub as well as the raw data, the scripts and the RStudio project [https://github.com/lschunk/edge_angle_analysis]. The analysis of the data aims at illustrating possible changes for instance along the tool edges and towards the surfaces.

USE-WEAR ANALYSIS

Stone tools enabled past humans to access food resources like animals and plants (Lombard 2005; McPherron et al. 2010; Wilkins et al. 2012). At the same time, they are the most commonly found physical evidence and the only continuous record to study the evolution of human behaviour (Key Proffitt/de la Torre 2020). Stone tools witness technological developments and innovations. Investigation how early hominins produced, used and curated their tools is thus a prerequisite for understanding the Pleistocene record. Stone tools can be observed in a great variety concerning their morphology. Nevertheless, in principle they all have something in common: they all have a certain function or combine several functions. All stone tools display the capacity to be used for a function, for instance modifying material, for movements e. g. scraping and cutting or for displaying a symbolic reason. Additionally, a tool's function can change over time or as a consequence of its use. Investigating the function of an artefact has long been part of archaeological research, either through the study of lithic typology and technology or directly through use-wear analysis. Use-wear analysis originates from the late nineteenth century, early twentieth century (Grace 1996). It includes the study of characteristic patterns of macroscopic and microscopic traces left on a tool, resulting from its use. Use-wear analysis is therefore the only approach providing a direct proof of a tool's utilisation (Shea 2011). Although the existence of retouch on a tool might imply tool use, solely use-wear analysis can confirm the actual use. The processes as to how use-wear forms are dependent on several aspects. For instance, the different types of use-wear traces respond to the raw material of the tool itself and the contact material, the tool morphology as well as the performed action and its duration (Kamminga 1979; Astruc/Vargiolu/Zahouani 2003; Lerner et al. 2007).

During the last couple of decades, use-wear analysis as a sub-discipline of archaeology passed through several changes, including methodological developments, theoretical and conceptual shifts. This process is also reflected by the given terminology. Discussions about the interchangeable character of these names and the definition of the discipline are thus the result (Marreiros et al. 2015).

Following Marreiros et al. (2020), the respective terminology – traceology, use-wear analysis and functional analysis – implies different levels of analysis and different objectives. To start with use-wear analysis, it defines the study of physical traces (microscopically and macroscopically) visible on an artefact's surface (Thomas et al. 2001; Marreiros et al. 2015). These traces are caused by human use. Traceology thereby goes a bit further. While it refers to the study of use-wear traces, it also includes traces resulting from production, non-utilitarian wear, and post-depositional traces. Ideally, use-wear analysis as well as traceology can and should be complemented with residue analysis. Functional analysis, however, combines methods such as technology, typology and use-wear analysis. Also, residue analysis can be part of functional analysis (Pe-

dergnana 2000). The aim of it has to be defined in a broader sense than only to identify the use of a tool. Functional studies target the evaluation of tool design, function and utilisation.

Traceology, use-wear analysis & functional analysis

In the following part, the history of this sub-discipline is summarised. Giving this context is of relevance, because it illustrates and justifies the steps and methods chosen and applied within this project. It also puts some emphasis on the circumstances as to why only a combination of different scales of analysis and methods can add more – and also new – information about *Keilmesser* and *Prądnik scrapers* to our current point of view.

The beginning of traceology/use-wear analysis can be traced back until the end of the nineteenth century. John Evans was presumably one of the first scientists stating the fact that the use of a tool can leave corresponding traces on its surface (Tringham et al. 1974). Equally early is the work of William Greenwill (Hayden 1979; Kamminga 1979). He identified use-wear traces on Palaeolithic tools and recognized the potential of the discipline. Nevertheless, it would take decades until traceology/use-wear analysis found a broader attention. This was thanks to Sergei Semenov (Semenov 1957; 1964), who created a methodological and systematic framework for the study of use-wear traces. For his work, Semenov is valued as the pioneer of the discipline. His publication *Prehistoric Technology* provides functional interpretations and experimental comparisons based on microscopic analysis. With his work, for the first time, experimental data has been applied as proxy for the identification of use-wear traces on archaeological samples. For his analysis, Semenov used a low magnification stereomicroscope (low-power approach) which led to a promotion of a systematic use of microscopic observations mainly on bones and lithics.

At the end of the 1960s, Semenov's work was translated and hence accessible to Western European scholars. Together with the rise of the *New Archaeology* (Binford 1962; Longacre 2010), it marked the beginning of a new period. The *New Archaeology* put the emphasis on developing scientific, hypothetic-deductive methods for the understanding of past human behaviour. The interest in for example ethnoarchaeology and the use of analogies grew. Traceology/use-wear analysis with its focus on tool function and utilisation and the microscopic approach fit the zeitgeist of the 1960s. Several researchers, for example Ruth Tringham (1974), George Odell (1975), Brian Hayden (1979) and Lawrence Keeley (1974; 1980), to mention only a few, concentrated on the new approach and tested its limitations. Consequently, the applied methods have been developed and improved, mainly by employing different microscopic magnifications. The end of the 1970s and the beginning of the 1980s are to mention as the timeframe for the introduction of the high-power approach (Keeley/Newcomer 1977). Compared to the low-power approach, the high-power approach is based on the use of high magnification microscopes, mostly upright metallographic microscopes. Since then, both approaches are used complementary, in the awareness of their limitations respectively. Additionally, the focus on experimental replications grew, achieving fundamental insights into the categorisation of diagnostic wear traces (Hardy/Garufi 1998; Vaughan 1985). With the aid of experimental samples, it could be demonstrated that the tool's raw material influences the nature of the resulting wear pattern. A comparison between archaeological and experimental assemblages requires therefore an inclusion of the corresponding raw material (Burroni et al. 2002; Evans/Donahue 2005).

The following years were characterised by improvements such as the development of more reliable and accurate methodologies by implementing more experiments and blind tests. At the same time, the limitations of a qualitative use-wear analysis became evident (Keeley 1974; Schiffer 1979; Grace/Graham/Newcomer 1985; Newcomer/Grace/Unger-Hamilton 1986; 1988; Moss 1987; Bamforth 1988; Hurcombe 1988).

These can be mainly summarised by the complexity of understanding use-wear traces, solely by the use of a stereomicroscope.
A new era was marked with the introduction of quantitative use-wear analysis. This significant step in the history of use-wear analysis was mainly possible due to the integration of a range of imaging equipment. The introduction of the following equipment and software is to name (Marreiros et al. 2020): the tactile profilometer (Beyries/Delamare/Quantin 1988), the atomic force microscopy (Kimball/Kimball/Allen 1995; Faulks et al. 2011; Kimball et al. 2017), the interferometry (Dumont 1982; d'Errico/Backwell 2009), the laser profilometry (Stemp/Stemp 2001; 2003), the confocal microscopy (Evans/Donahue 2008; Stemp/Chung 2011; Stemp/Lerner/Kristant 2013; Evans 2014; Macdonald/Harman/Evans 2018), and the image analysis and surface metrology software (d'Errico/Backwell 2009; Sahle et al. 2013; Ibáñez/Lazuen/González-Urquijo 2018; Martisius et al. 2018; 2020; Calandra et al. 2019a, b). The aim of quantitative use-wear analysis is the identification of different use-wear traces following standardised, quantitative criteria based on 2D and 3D images and surface roughness measurements. More recently, new attempts to quantify surface textures have been made (Evans/Macdonald 2011; Stemp et al. 2015; Martisius et al. 2018; Ibáñez/Lazuen/González-Urquijo 2018; Pedergnana et al. 2020b; Pedergnana/Ollé/Evans 2020). Still, criticism has been raised against a lack of reproducibility regarding the acquired data (Evans et al. 2014; Calandra et al. 2019c; Marreiros et al. 2020). While traceology/use wear analysis as a sub-discipline of archaeology is still continuously improving, scientists are aware of the current limitations of the approach. The integration of established methods from related disciplines (e. g. dental microwear) could offer great potential for further improvements (Calandra et al. 2019b).

Methods and techniques in use-wear analysis

Initial use-wear studies were characterised by the use of the low-power approach. Since the introduction of a systematic use of low magnification stereomicroscopy by Semenov, this was the state-of-the-art for use-wear analysis. This changed in the 1970s, when Keeley introduced the high-power approach. He used a reflected light microscope for his work. After years of discussion and debates about the advantages and disadvantages of both approaches, a broad agreement has been reached. Only a complementary method by the combination of low- and high-power approach can lead to a profound analysis and a reliable result (Odell 2001). The use of the low-power approach should thereby allow the identification of macro-wear traces as for example edge damage and impact fractures. Moreover, this approach should help with identifying the areas of interest for the subsequent high-power approach. The observation under higher magnifications should allow for a detailed analysis of micro-wear traces as for instance striations and polish formations (Marreiros et al. 2015).
In general, the methods applied in use-wear analysis nowadays can be distinguished in digital microscopy, II) optical microscopy (stereomicroscopes, upright metallographic microscopes), III) scanning electron microscopy (SEM), and IV) laser scanning confocal microscopy (LSM).
When referring to the low-power approach, a stereomicroscopic analysis in the range of four to ten × magnification is common. Enabled by a movable light source, reflected light, the artefact can be illuminated from different angles, allowing for a shadow effect and thus an easier detection of possible traces. The implementation of the high-power approach is usually characterised by a metallographic microscope and magnifications between 50 and 400×. An incident light serves as a light source perpendicular (90°) to the material surface.

Another high magnification observation method offers the scanning electron microscopy (SEM). An SEM uses instead of a light illumination a focused beam of electrons controlled by magnetic or electric fields. The advantages of an SEM are the higher possible magnifications, resolutions, and depth of field compared to a metallographic microscope (Hay 1977; Del Bene 1979). The use of an SEM has also proven to lead to better results with certain raw materials, for example quartzite (Ollé et al. 2016; Pedergnana 2017). However, the use of an SEM is also accompanied with disadvantages. Commonly, an SEM provides only a small chamber for the sample, meaning the object volume is limited. A sample preparation is typically needed, sometimes including a sample coating. Furthermore, an SEM is a costly piece of equipment and the analysis is relatively time-consuming. These aspects taken together usually lead to an analysis of a comparable small sample size. Studies have demonstrated the effectiveness of a complementary approach, the combination of optical and electron microscopy (Monnier/Ladwig/Porter 2012; Borel et al. 2014; Ollé et al. 2016).
A quantitative use-wear analysis can be done by the means of laser scanning microscopy. The laser scanning microscopy (LSM) is an optical imaging technique. The illumination of the sample happens through a spatial pinhole system, which blocks out-of-focus light. This means, only one point on the sample at a time is illuminated. In order to acquire 2D or 3D images, the sample has to be scanned over a raster. An LSM allows observations ranging between 25-800× magnification (e. g. Mansur-Franchomme 1983; Derndarsky/Ocklind 2001; Shanks et al. 2001; Scott et al. 2005; 2006; Debert/Sherriff 2007; Evans/Donahue 2008). Laser confocal microscopy has been used mainly in archaeology for the illustration and the modelling of surface topography. Several parameters, also in accordance with ISO Norms, can be calculated on the tool's surface (e. g. amplitude parameters, spatial parameters, roughness parameters) (Calandra et al. 2019c). Based on basic roughness parameters it is possible to distinguish diagnostic use-wear traces (Giusca et al. 2012). During the last years, the LSM as well as the SEM have been used more frequently in use-wear analysis (Stemp/Stemp 2001; 2003; Lerner et al. 2007; Evans/Donahue 2008; Evans/Macdonald 2011; Stemp/Chung 2011; Giusca et al. 2012; Pedergnana et al. 2020b; Pedergnana/Ollé/Evans 2020). The likely explanation for this is the possibility to quantify micro-wear traces with both types of equipment. Unfortunately, the LSM is an expensive purchase, but the running costs are smaller and the limitations regarding the sample size are less restricting.

Artefact cleaning procedure

Before doing the use-wear analysis, the sampled artefacts had to be cleaned (Pedergnana et al. 2016; Pedergnana et al. 2020a). To do so, each sample was individually packed in a plastic bag filled with ~ 100 ml of demineralised water and a non-ionic detergent (BASF Plurfac LF901, 1 g/l = 1 % w/v; BASF SE, Ludwigshafen, Germany). The closed bags were put into a preheated ultrasonic bath (EMAG Emmi 20HC). The samples were left in the bath for 15 min at 40°C and 100 KHz.
Thereafter, the samples, still placed in the bags, were rinsed three times with tap water. This step is meant to remove the surfactant residues. The bags were then filled once with ~100 ml purified water and emptied after rinsing. After that, the samples were air-dried. Immediately before and during the data acquisition, the measured surface was cleaned with 2-propanol 70 % v/v.

Qualitative use-wear analysis

The methodological approach of use-wear analysis can be divided in a low-power and high-power approach (Keeley 1974; 1980; Marreiros et al. 2015). While the low-power approach is equal to a macroscopic meth-

odology, the high-power approach involves the use of optical microscopy magnification. The sampling for the use-wear analysis was done on a macroscopic observation level. When necessary, the artefacts were observed regarding their preservation with a stereomicroscope (ZEISS SteREO Discovery.V8). Only complete lithics representing all existing categories as well as *Keilmesser* shapes from the two different raw materials were sampled.

The qualitative use-wear analysis was done in a high-power approach by means of an upright light microscope (ZEISS Axio Scope.A1 MAT; **tab. 10**). The samples were studied with a 5×, 10× and 20× magnification. Traces were documented as an EDF image in black and white settings.

Tool areas displaying use-wear traces were afterwards documented with a digital microscope (ZEISS Smartzoom 5) by the use of a 1.6× objective and 34-× zoom. The combination of stitching and EDF makes a documentation of an entire tool edge in focus possible.

		LSM	upright light microscope	digital microscope	3D scanner
Microscope	Manufacturer	Carl Zeiss Microscopy GmbH	Carl Zeiss Microscopy GmbH	Carl Zeiss Microscopy GmbH	AICON, now part of Hexagon AB
	Model	Axio Imager.Z2 Vario + LSM 800 MAT	Axio Scope.A1 MAT	Smartzoom 5	smartScan-HE R8
Location	Laboratory	TraCEr, MONREPOS, Germany	TraCEr, MONREPOS, Germany	TraCEr, MONREPOS, Germany	TraCEr, MONREPOS, Germany
	Floor	basement (-1)	basement (-1)	basement (-1)	basement (-1)
	Setup	Passive anti-vibration table on solid concrete base	Standard office desk	Stable table on solid concrete base	column stand
Acquisition	Software	ZEN blue 2.3 with Shuttle&Find module	Zen core 2.7	Smartzoom software with Shuttle&Find module	OptoCat 2018R1
	Mode	LSM (laser scanning confocal microscopy)	Bright field	Bright field	Structured light
Objective	Manufacturer	Carl Zeiss Microscopy GmbH	Carl Zeiss Microscopy GmbH	Carl Zeiss Microscopy GmbH	Hexagon
	Objective	C Epiplan-Apochromat 10× / NA = 0.40 / WD = 5.4 mm	EC Epiplan 5× / NA = 0.13 / WD = 11.8 mm	PlanApo D 1.6× / NA = 0.1 / WD =36 mm	M-450, 108 µm point-to-point distance
		C Epiplan-Apochromat 20× / NA = 0.70 / WD = 1.30 mm	EC Epiplan 10× / NA = 0.25 / WD = 11.0 mm	-	S-150, 33 µm point-to-point distance
		C Epiplan-Apochromat 50× / NA = 0.95 / WD = 0.22 mm	EC Epiplan 20× / NA = 0.40 / WD = 3.2 mm	-	-

Table 10 Acquisition settings for the involved pieces of equipment including the objective specifications and the resulting magnifications and resolutions. NA = numerical aperture, WD = working distance, FOV = field of view. The total on-screen magnification can be calculated as follows: objective magnification × optical zoom × camera adaptor × screen diagonal × digital zoom camera/camera sensor diagonal.

		LSM	upright light microscope	digital microscope	3D scanner
Illumination	Source	Laser	White LED, reflected co-axial light **	White LED, reflected ring-light	Blue LED
	Wavelength	405 nm	550 nm	550 nm	-
Settings	Master Gain	245 V	-	-	-
	Scanning direction	Both ways (no correction, line step = 1)	-	-	-
	Scanning speed	8 (max)	-	-	-
	Bit depth	16 bits	-	-	-
	Pinhole diameter	54 µm (50× obj.) / 34 µm (20× obj.) (1 AU lateral optical resolution)	-	-	-
Size and resolution	Zoom	0.5× (50× obj.) / 0.2× (20× obj.)	-	-	-
	Step size	0.25 µm	-	-	-
	Data quality	No noise cut (0-65335 levels, post-processing)	-	-	-
	FOV	255.56 × 255.56 µm (50× obj.) / 638.9 × 638.9 µm (20× obj.)	1.70 × 1.42 mm (5× obj.) / 850.08 × 709.32 µm (10× obj.) / 425.04 × 354.66 µm (20× obj.)	10.475 7.856 mm	355 × 265 × 220 mm measuring volume (M-450) / 110 × 80 × 70 mm measuring volume (S-150)
	Frame size	1198 × 1198 pi×els (50× obj.) / 3000 × 3000 pi×els (20× obj.)	2464 × 2056 pi×els	15875 × 3301 pixels	-
	Total on-screen magnification	1318× (50× obj.) / 527× (20× obj.)	187× (5× obj.) / 375× (10× obj.) / 750× (20× obj.)	34×	-

Table 10 (continued)

Spatial pattern recognition

A scheme was developed to locate and map use-wear traces (for similar scheme see Plisson 1985; Van Gijn 1990; Lombard 2008; Mazzucco 2018). The outline of the artefact was used as a grid (**fig. 26**). *Keilmesser*, *Prądnik scrapers*, scraper and flakes were separated in a grid of six areas, *Prądnik spalls* in four areas. This was done for the dorsal (A and B) as well as for the ventral (D and C) surface. The areas are numbered from one, distal part of the tool, to two, the medial part, and three, the proximal part of the tool. Following this scheme, use-wear traces can easily be located and a map for all use-wear traces can be generated, allowing for statements regarding the frequency and distribution of the use-wear traces. Thus, this qualitative use-wear approach aims for pattern recognition (e. g. location of the use-wear traces on tool's surface). Included in the scheme is the type of use-wear and the orientation. Possible predefined use-wear types are polish, striations and impact marks. Here, polish refers to a modification of the tool surface caused by contact between the tool and the worked material. Polish appears as a dull, brighter or smoothed area on the surface (Keeley 1980; Haslam 2009; Rots 2013; Evans 2014). The term striation refers to linear scratches, furrows or grooves

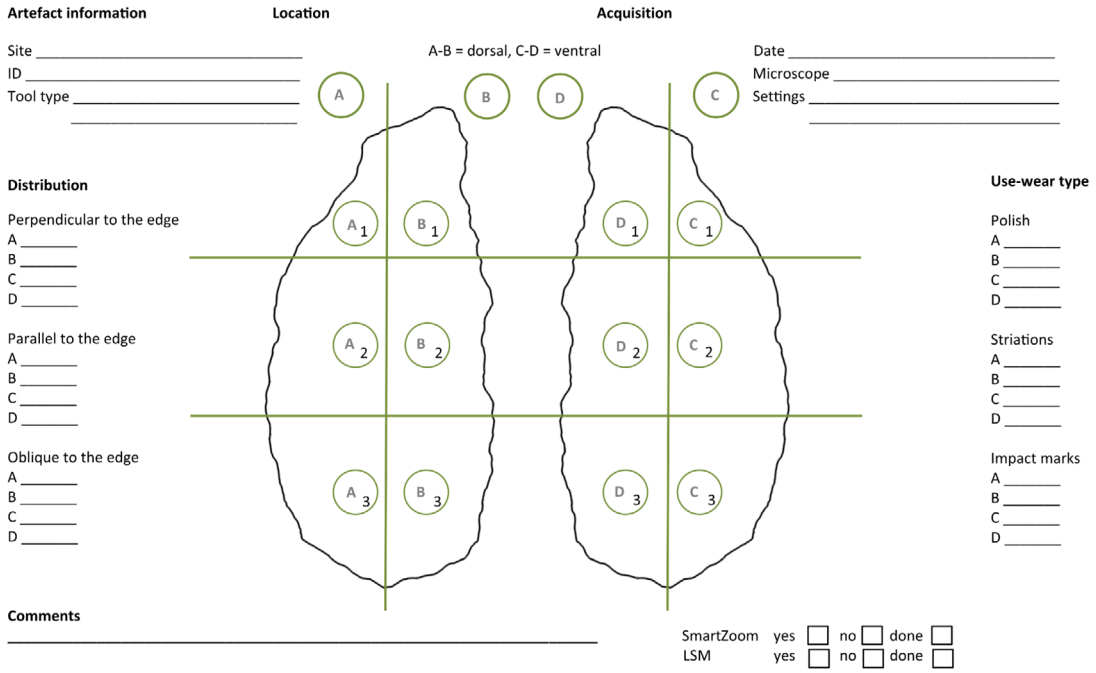


Fig. 26 Scheme used during the qualitative use-wear analysis for the documentation of the analysed tools. Here, the outline represents a *Keilmesser*. The dorsal (A and B) and ventral surfaces (D and C) are exemplarily labelled as if the tool would be a right-sided *Keilmesser*. However, the outline is different for the other artefact categories. Additionally, the scheme includes information about the analysed artefact itself, the acquisition date and settings, as well as information about the use-wear traces (location, distribution and type). The scheme also contains information, whether further acquisition or analysis are needed (SmartZoom = digital microscope images, LSM = quantitative use-wear analysis).

reflecting an abrasive action between the tool and the worked material (Keeley 1980). Moreover, the orientation for the use-wear traces can be specified with the options perpendicular, parallel or oblique to the edge.

Description use-wear traces

The aim of use-wear analysis is to identify traces left on the tool's surface. The first step is the recognition and the characterisation of different types of traces, for example polish, impact fractures or striations. It is known that wear patterns correlate with the worked material (Keeley 1980; Haslam 2009; Rots 2013). A further step usually seeks, especially concerning polish, a more detailed description and identification of the observed traces. In the qualitative approach, this is exclusively based on visual identification of the traces. Polish is a modified area on the tool, recognisable with an optical microscope by a brighter or smoother appearance (Keeley 1980). It is caused by a gradual removal or deformation of the natural surface. Various names for different types of polish can be found in literature for example wavy polish, and also their functional interpretation as wood polish, bone polish, hide polish etc. Although the classification of polish and its formation is one of the fundamental questions addressed in use-wear analysis (Marreiros et al. 2015), a consensus and an agreement on the various terms and identifications is still missing (Grace/Graham/Newcomer 1985; Grace 1989; Evans 2014; Van Gijn 2014; Marreiros et al. 2020). The different types of polish are likely related to several other aspects such as the raw material, the use intensity, the performed task and not only the worked material (Buc 2011). This means, even the same combination of tool and worked material can lead to varying traces by performing the same movement when only the duration differs. This makes it extremely difficult to assess the exact use of a tool confidentially by visual means only.

Here, documented traces were visually inspected and categorised. A subjective description of these categorised wear traces was done. With respect to the various existing definitions, nine types of use-wear traces were defined for the samples from the three analysed assemblages. Sometimes, the distinction between two types of use-wear traces was ambiguous. In these cases, the combination of two of the defined types was chosen.

Analysis use-wear data

After documenting and defining the use-wear traces on all analysed artefacts, the manually filled scheme for the use-wear pattern recognition (see section 4.4.3.1) had to be digitalised. Therefore, the initially used artefact outline templates (SVG files) were converted into DXF format using the open-source software Potrace tracing program (version 1.1.6), producing internal and external outlines as well as the grid. The DXF were modified using CAD. The resulting DXF files were then imported into QGIS (version 3.14.16) templates as vectors. QGIS is a free and open source geographic information system, which was used here to map the documented traces. Next to the described vector layers, additional point layers for each individual type of use-wear trace were created. The point layers were given a suitable colour scheme. This method allows a simple and visual presentation of all use-wear traces separated per artefact type.

Quantitative use-wear analysis

Based on the results from the previously described qualitative use-wear analysis, samples were selected for a quantitative use-wear analysis. Thus, the quantitative use-wear analysis was done on samples displaying different types of use-wear and use-wear on the different locations. The minimum number of measurements per use-wear type was three. The 3D surface topography data was acquired with a confocal microscope. The confocal microscope is an upright light microscope coupled with a laser-scanning confocal microscope (ZEISS Axio Imager.Z2 Vario + ZEISS LSM 800 MAT; **tab. 10**).
When time played a limiting factor, for example during a short time loan from a museum or during the execution of an experiment, moulds of the region of interest were taken. The moulds assure a possibility of a quantitative use-wear analysis at a later time, since they serve as an identical copy of the real surface. The moulds were made of Provil novo Light regular (Kulzer GmbH, Hanau, Germany), a two-component silicone impression material applied with a dispensing gun. Under regular room conditions, the moulds are fully hardened after 5-8 minutes.
In total 50 different use-wear spots were measured with the LSM. Six of these use-wear traces are on artefacts from Buhlen, 33 from Balver Höhle and eleven from Ramioul. Four of these 50 measurements were taken on moulds instead of the original artefact surface.
Before using the LSM, the system always had to be turned on one hour in advance in order to warm up all components and to limit thermic drift. In order to assure that the artefacts cannot move and stay stable during the acquisition, a mould with a flat lower surface in the sense of a sample holder was made. The mould was made of Provil novo Putty regular set (Kulzer GmbH, Hanau, Germany). This set is a two-component silicone impression material consisting of a base paste and a matching catalyst paste. After combining these two components in a ratio 50:50, the curing time is roughly three to five minutes in room conditions. Additionally, a goniometer stage was used to position the sample with the area to measure as horizontal as possible. The LSM is equipped with a motorized revolver allowing an easy change of objectives. For the

acquisition with the LSM the aim was to scan the surface under the highest magnification, respectively. Therefore, the objective C Epiplan-Apochromat 50× with the numerical apertures 0.95 was used. The field of view (FOV) was 255.6 × 255.6 µm. Sometimes the artefact surfaces were not straight enough or the artefacts could not be positioned in a way needed in order to measure the surfaces with the 50× /0.95 objective, so that the C Epiplan 50× /0.75 (for nine samples) or the C Epiplan-Apochromat 20× /0.70 (for one sample) had to be used. While the 50× /0.75 has the same FOV than the other 50× objective, the 20× /0.70 objective has a FOV of 638.9 × 638.9 µm. Each use-wear trace was measured three times at nearby but non-identical spots. These scans are treated as replicas and verify (or not) the homogeneity within each trace.
Wide field images as an overview were not acquired with the LSM since EDF pictures were already taken with the upright light microscope.

Data processing

The acquired data was processed in batch for each archaeological site respectively with templates in ConfoMap (a derivative of MountainsMap Imaging Topography developed by Digital Surf, Besançon, France; version ST 8.1.9286; for details see https://guide.digitalsurf.com/en/guide.html). In total four different templates had to be applied. The first template – the »extracting template« – was only used for the 3D surfaces acquired with the 20× /0.70 objective (**figs 27-34**). After applying this template, all surfaces had the same size of 254.9 × 254.9 µm. The following second template, »resampling template« was by contrast used to perform a resampling in x and y on all 150 surfaces, leading to an identical spacing. For the scans acquired on moulds, an additional third template was needed, the »mirroring template«. Moulds always just display a mirrored version of the original surfaces. Therefore, the acquired surface needed to be mirrored in order to match the original surface again. In a final step, the data was processed with a »processing template«. The following procedure was performed: I.) Levelling (least squares method by subtraction), II.) From removal (polynomial of degree 3), III.) Outliers removal (maximum slope of 80°), IV.) Thresholding the surface between 0.1 and 99.9 % material ratio to remove the aberrant positive and negative spikes, V.) Applying a robust Gaussian low-pass S-filter (S_1 nesting index = 0.425 µm, corresponding to about five pixels, end effects managed) to remove noise, VI.) Filling-in the non-measured points (NMP), VII.) Analysis: Calculation of 21 ISO 25178-2 parameters, 3 furrow parameters, 3 texture direction parameters, one texture isotropy parameter and the scale-sensitive fractal analysis (SSFA) parameters *epLsar, NewEplsar, Asfc, Smfc, HAsfc9 and HAsfc81* (**tab. 11**).
The ConfoMap templates for each site and surface in MNT and PDF formats are available on GitHub [https://github.com/lschunk/Archaeology_use-wear]. This also includes all original and processed surfaces, as well as the results.

Statistical procedure

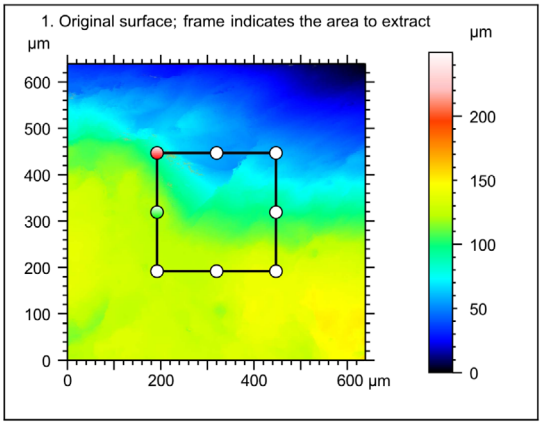
All descriptive analyses (summary statistics, scatter plots, box plots, histogram plots and principal component analysis, PCA) were performed in the open-source software R version 4.0.2 through RStudio version 1.3.1073 (RStudio Inc., Boston, USA) for Microsoft Windows 10. The following packages were used openxlsx v. 4.1.5, R.utils v. 2.9.2, doBy v. 4.6.7, ggrepel v. 0.8.2, patchwork v. 1.0.1, ggplot2 v. 3.3.2, tidyverse v. 1.3.0, wesanderson v. 0.3.6. Reports of the analysis in HTLM format, created with knitr v. 1.29 and rmarkdown v. 2.3 are available on GitHub [https://github.com/lschunk/Archaeology_use-wear]. Also the raw data, the scripts and the RStudio project are saved in the same repository.

Quantitative use-wear analysis: Balver Höhle
Template to extract a 255 x 255 µm surface acquired with the LSM800 with the 20x/0.7 objective in order to match the FOV of the 50x0.95 and 50x0.7 objectives

This template is used for one sample with three measurements (surfaces).

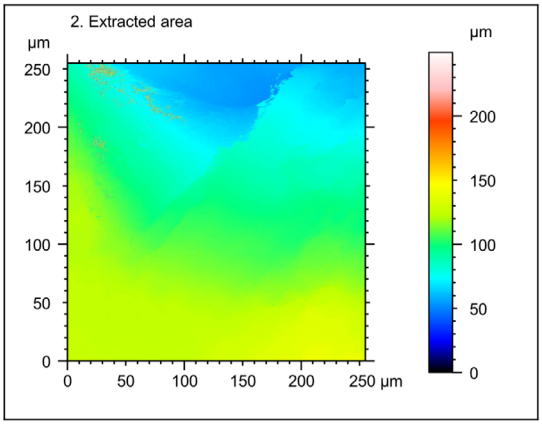
A. Extracting

1. Original surface; frame indicates the area to extract



Identity card		
Name:	MU-232-B2-01-a_20x07_LSM_Topo	
Created on:	7/7/2020 4:58:00 PM	
Studiable type:	Surface	
Axis:	X	
Length:	638.7	µm
Size:	3000	points
Spacing:	0.2130	µm
Axis:	Y	
Length:	638.7	µm
Size:	3000	points
Spacing:	0.2130	µm
Axis:	Z	
Layer type:	Topography	
Length:	249.7	µm
Size:	65531	digits
Spacing:	3.810	nm
NM-points ratio:	0.000 % (0 Pts)	

2. Extracted area



Identity card		
Name:	MU-232-B2-01-a_20x07…_Topo > Extracted area	
Created on:	7/7/2020 4:58:00 PM	
Studiable type:	Surface	
Axis:	X	
Length:	254.9	µm
Size:	1198	points
Spacing:	0.2130	µm
Axis:	Y	
Length:	254.9	µm
Size:	1198	points
Spacing:	0.2130	µm
Axis:	Z	
Layer type:	Topography	
Length:	249.6	µm
Size:	65505	digits
Spacing:	3.810	nm
NM-points ratio:	0.000 % (0 Pts)	

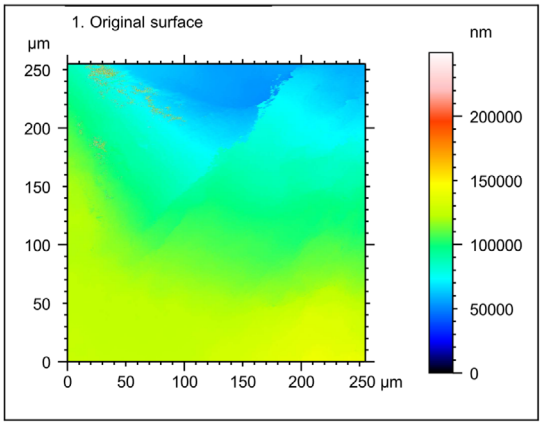
ConfoMap ST 8.1.9286 23-Jul-20 - 1 / 1 -

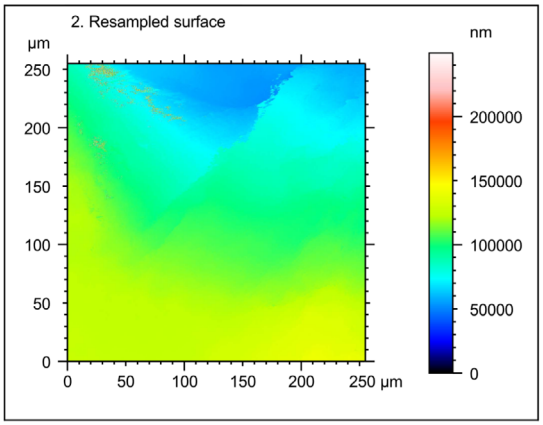
Fig. 27 »Extracting template« created in ConfoMap (a derivative of MountainsMap Imaging Topography developed by Digital Surf, Besançon, France; version ST 8.1.9286) to process the data acquired from the quantitative use-wear analysis. Here, the template was applied to a *Keilmesser* from Balver Höhle (MU-232). The template was only used for samples analysed with the 20×/0.7 objective.

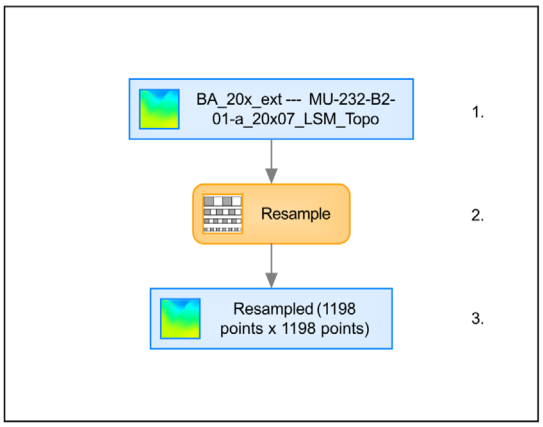
Quantitative use-wear analysis: Balver Höhle
Template to resample in x and y in order to reach an identical spacing in all with the LSM800 acquired surfaces (50x/0.7 and 50x/0.95 objectives).

This template is used for thirty-three samples with three measurements (surfaces) respectively.

B. Resampling







Identity card		
Name:	BA_20x_ext --- MU-232…01-a_20x07_LSM_Topo	
Created on:	7/7/2020 4:58:00 PM	
Studiable type:	Surface	
Axis:	X	
Length:	254.9	µm
Size:	1198	points
Spacing:	0.2130	µm
Axis:	Y	
Length:	254.9	µm
Size:	1198	points
Spacing:	0.2130	µm
Axis:	Z	
Layer type:	Topography	
Length:	249564	nm
Size:	65505	digits
Spacing:	3.810	nm
NM-points ratio:	0.000 % (0 Pts)	

Identity card		
Name:	BA_20x_ext --- MU-232…8points x 1198 points)	
Created on:	7/7/2020 4:58:00 PM	
Studiable type:	Surface	
Axis:	X	
Length:	254.9	µm
Size:	1198	points
Spacing:	0.2130	µm
Axis:	Y	
Length:	254.9	µm
Size:	1198	points
Spacing:	0.2130	µm
Axis:	Z	
Layer type:	Topography	
Length:	249564	nm
Size:	65505	digits
Spacing:	3.810	nm
NM-points ratio:	0.000 % (0 Pts)	

ConfoMap ST 8.1.9286

25-Aug-20

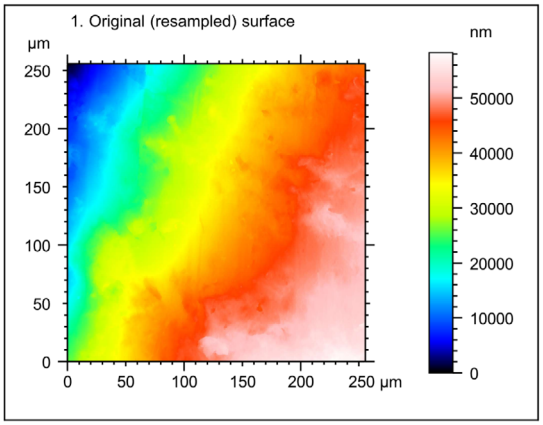
- 1 / 1 -

Fig. 28 »Resampling template« created in ConfoMap (a derivative of MountainsMap Imaging Topography developed by Digital Surf, Besançon, France; version ST 8.1.9286) to process the data acquired from the quantitative use-wear analysis. Here, the template was applied to a *Keilmesser* from Balver Höhle (MU-232).

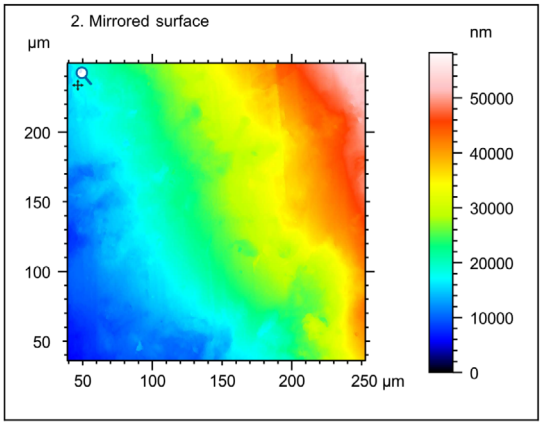
Quantitative use-wear analysis: Balver Höhle
Template to mirror the surfaces in x and z that have been acquired with the LSM800 based on moulds instead of the original artefact surface.

This template is used for three samples with three measurements (surfaces) respectively.

C. Mirroring



Identity card		
Name:	BA_50x_res --- MU-020 …01-a_50x09_LSM_Topo	
Created on:	7/8/2020 4:41:00 PM	
Studiable type:	Surface	
Axis:	X	
Length:	255.5	µm
Size:	1198	points
Spacing:	0.2134	µm
Axis:	Y	
Length:	255.5	µm
Size:	1198	points
Spacing:	0.2134	µm
Axis:	Z	
Layer type:	Topography	
Length:	58262	nm
Size:	65509	digits
Spacing:	0.8894	nm
NM-points ratio:	0.000 % (0 Pts)	



Identity card		
Name:	BA_50x_res --- MU-020 …> Mirrored (in X and Z)	
Created on:	7/8/2020 4:41:00 PM	
Studiable type:	Surface	
Axis:	X	
Length:	255.5	µm
Size:	1198	points
Spacing:	0.2134	µm
Axis:	Y	
Length:	255.5	µm
Size:	1198	points
Spacing:	0.2134	µm
Axis:	Z	
Layer type:	Topography	
Length:	58262	nm
Size:	65509	digits
Spacing:	0.8894	nm
NM-points ratio:	0.000 % (0 Pts)	

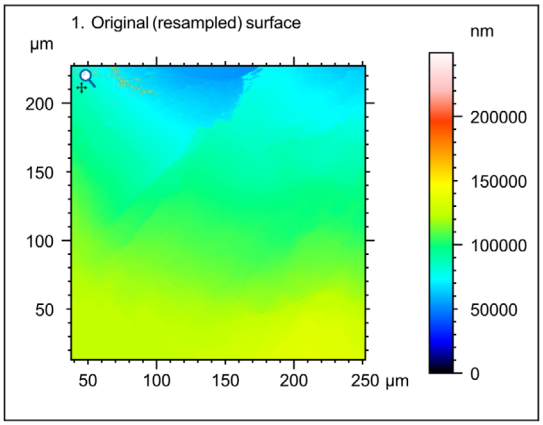
ConfoMap ST 8.1.9286 25-Aug-20 - 1 / 1 -

Fig. 29 »Mirroring template« created in ConfoMap (a derivative of MountainsMap Imaging Topography developed by Digital Surf, Besançon, France; version ST 8.1.9286) to process the data acquired from the quantitative use-wear analysis. Here, the template was applied to a *Keilmesser* from Balver Höhle (MU-020). This template was only used when a mould instead of the original artefact surface was analysed.

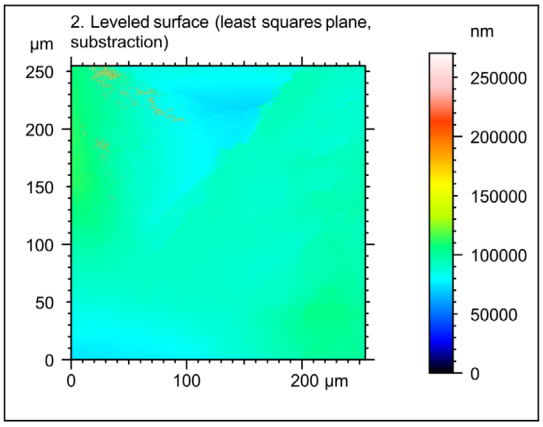
Quantitative use-wear analysis: Balver Höhle
Template to process all surfaces aquired with the LSM with the 20x/0.7, 50x/0.75 and 50x/0.95 objectives.

This template is used for thirty-three sampleswith three measurements(surfaces) respectively.

D. Processing



Identity card		
Name:	BA_50x_res --- BA_20x...01-a_20x07_LSM_Topo	
Created on:	7/7/2020 4:58:00 PM	
Studiable type:	Surface	
Axis:	X	
Length:	254.9	µm
Size:	1198	points
Spacing:	0.2130	µm
Axis:	Y	
Length:	254.9	µm
Size:	1198	points
Spacing:	0.2130	µm
Axis:	Z	
Layer type:	Topography	
Length:	249564	nm
Size:	65505	digits
Spacing:	3.810	nm
NM-points ratio:	0.000 % (0 Pts)	



Identity card		
Name:	BA_50x_res --- BA_20x...po> Leveled (LS-plane)	
Created on:	7/7/2020 4:58:00 PM	
Studiable type:	Surface	
Axis:	X	
Length:	254.9	µm
Size:	1198	points
Spacing:	0.2130	µm
Axis:	Y	
Length:	254.9	µm
Size:	1198	points
Spacing:	0.2130	µm
Axis:	Z	
Layer type:	Topography	
Length:	270491	nm
Size:	70998	digits
Spacing:	3.810	nm
NM-points ratio:	0.000 % (0 Pts)	

ConfoMap ST 8.1.9369 07-Sep-20 - 1 / 5 -

Fig. 30 »Processing template« created in ConfoMap (a derivative of MountainsMap Imaging Topography developed by Digital Surf, Besançon, France; version ST 8.1.9286) to process the data acquired from the quantitative use-wear analysis. Here, the template was applied to a *Keilmesser* from Balver Höhle (MU-232). This is page 1 of 5.

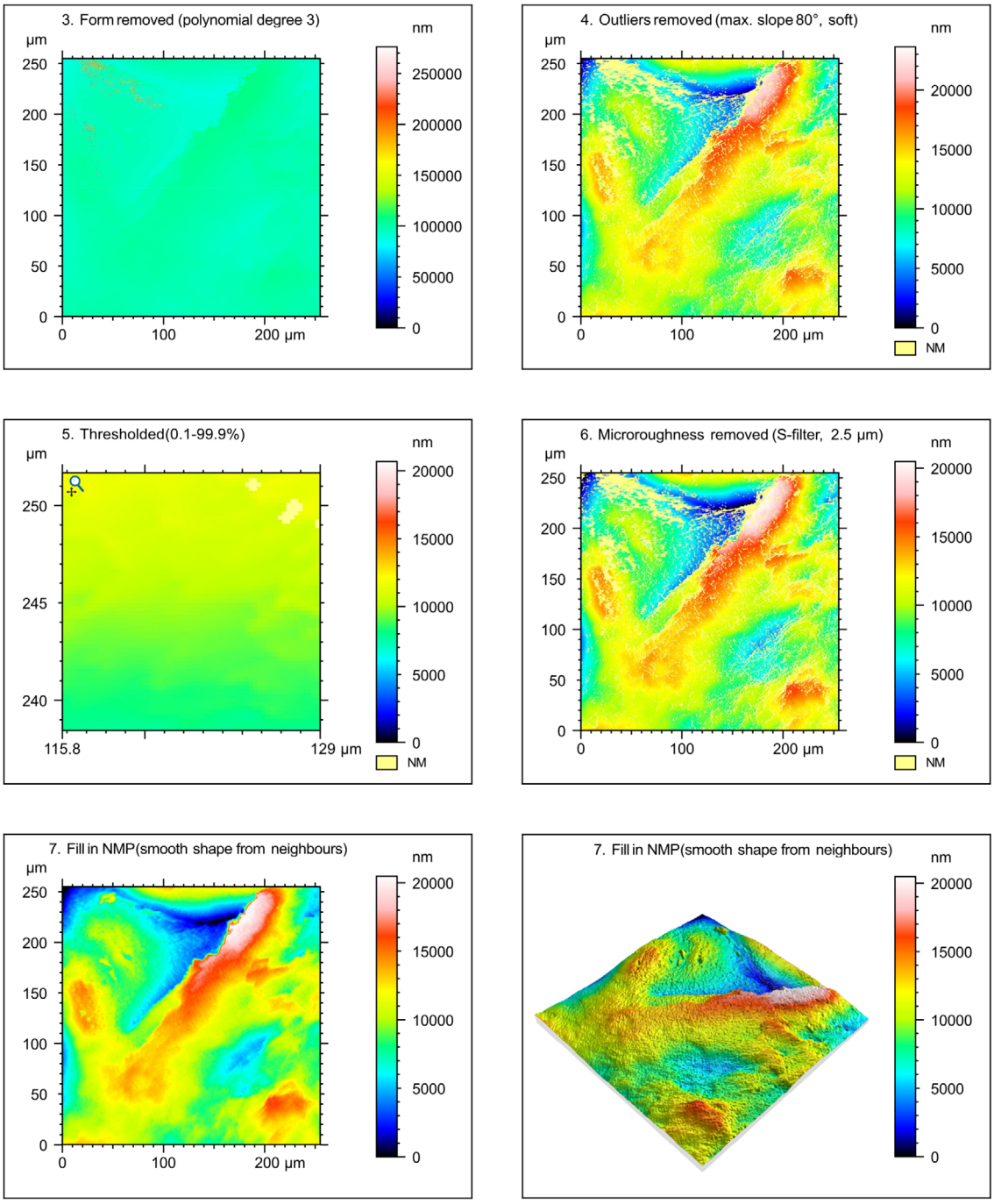


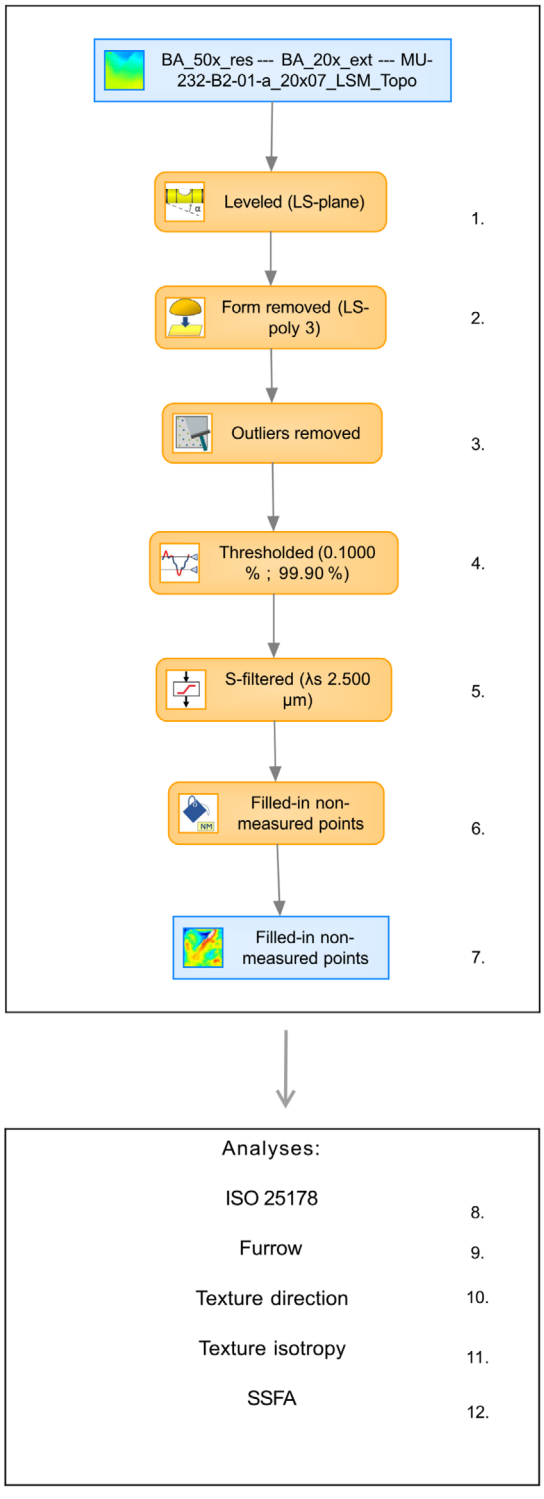
Fig. 31 »Processing template« created in ConfoMap (a derivative of MountainsMap Imaging Topography developed by Digital Surf, Besançon, France; version ST 8.1.9286) to process the data acquired from the quantitative use-wear analysis. Here, the template was applied to a *Keilmesser* from Balver Höhle (MU-232). This is page 2 of 5.

Identity card			
Name:	BA_50x_res --- BA_20x...innon-measured points		
Created on:	7/7/2020 4:58:00 PM		
Studiable type:	Surface		
Axis:	X		
Length:	254.9	µm	
Size:	1198	points	
Spacing:	0.2130	µm	
Axis:	Y		
Length:	254.9	µm	
Size:	1198	points	
Spacing:	0.2130	µm	
Axis:	Z		
Layer type:	Topography		
Length:	20482	nm	
Size:	53762	digits	
Spacing:	0.3810	nm	
NM-points ratio:	0.000 % (0 Pts)		

C. Analyses

8. ISO 25178-2 parameters on surface #7

ISO 25178 - Primary surface		
F: [Workflow] Form removed (LS-poly 3)		
S-filter (λs): [Workflow] S-filtered (λs 2.500 µm)		
Height parameters		
Sq	3243	nm
Ssk	0.06336	
Sku	3.457	
Sp	10477	nm
Sv	10005	nm
Sz	20482	nm
Sa	2506	nm
Functional parameters		
Smr	0.5510	%
Smc	3754	nm
Sxp	6582	nm
Spatial parameters		
Sal	25.95	µm
Str	0.3211	
Std	42.50	°
Hybrid parameters		
Sdq	0.6025	
Sdr	9.994	%
Functional parameters (Volume)		
Vm	0.2094	µm³/µm²
Vv	3.963	µm³/µm²
Vmp	0.2094	µm³/µm²
Vmc	2.775	µm³/µm²
Vvc	3.559	µm³/µm²
Vvv	0.4034	µm³/µm²



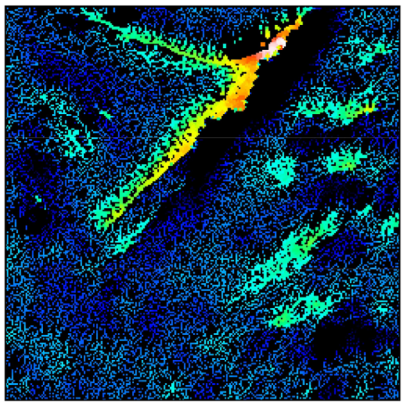
ConfoMap ST 8.1.9369

07-Sep-20

- 3 / 5 -

Fig. 32 »Processing template« created in ConfoMap (a derivative of MountainsMap Imaging Topography developed by Digital Surf, Besançon, France; version ST 8.1.9286) to process the data acquired from the quantitative use-wear analysis. Here, the template was applied to a *Keilmesser* from Balver Höhle (MU-232). This is page 3 of 5.

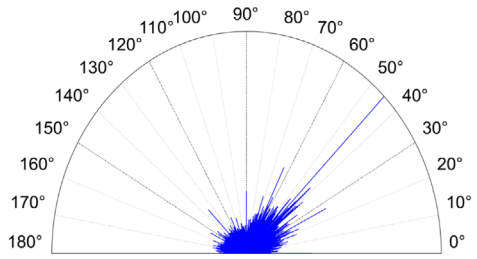
9. Furrow analysis on surface #7



All furrows are shown.

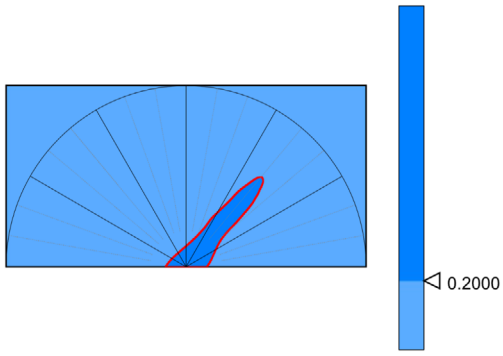
Parameters	Value	Unit
Maximum depth of furrows	12698	nm
Mean depth of furrows	2586	nm
Mean density of furrows	2987	cm/cm2

10. Texture direction on surface #7



Parameters	Value	Unit
First direction	44.98	°
Second direction	26.45	°
Third direction	63.53	°

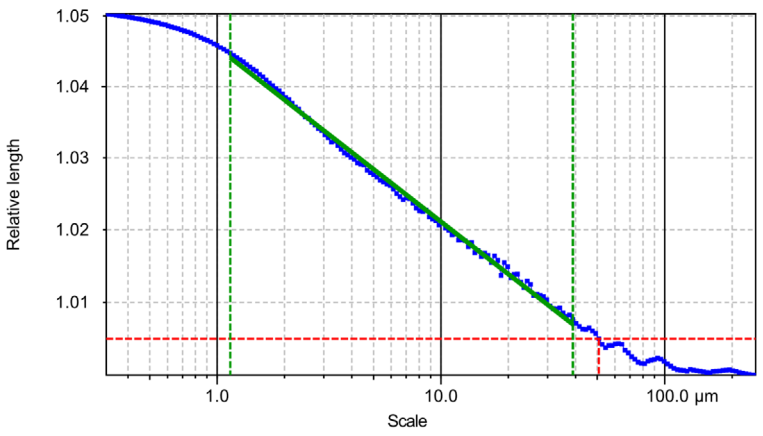
11. Texture isotropy on surface #7



Parameters	Value	Unit
Isotropy	13.50	%

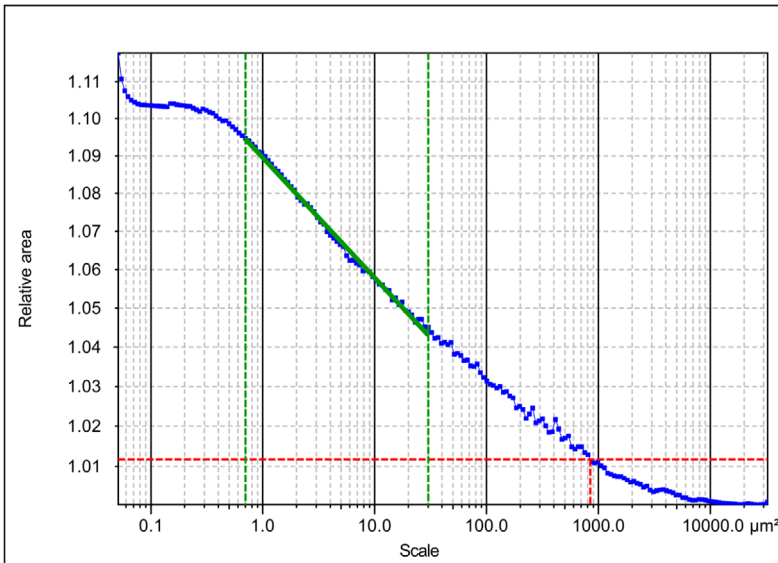
Fig. 33 »Processing template« created in ConfoMap (a derivative of MountainsMap Imaging Topography developed by Digital Surf, Besançon, France; version ST 8.1.9286) to process the data acquired from the quantitative use-wear analysis. Here, the template was applied to a *Keilmesser* from Balver Höhle (MU-232). This is page 4 of 5.

12. SSFAon surface #7



Information	
Method	Length-scale (rows)

Parameters	Value	Unit	Comment
epLsar	0.003683		*Length-scale anisotropy (Sfrax) (1.8 µm, 5°)*
NewEplsar	0.0181		*Length-scale anisotropy (1.8 µm, 5°)*



Information	
Method	Area-scale (four corners)

Parameters	Value	Unit	Comment
Asfc	12.80		*Fractal complexity*
Smfc	2.508	µm²	*Scale of max complexity*
HAsfc9	0.6292		*Heterogeneity of Asfc (3x3)*
HAsfc81	0.8102		*Heterogeneity of Asfc (9x9)*

ConfoMap ST 8.1.9369

07-Sep-20

- 5 / 5 -

Fig. 34 »Processing template« created in ConfoMap (a derivative of MountainsMap Imaging Topography developed by Digital Surf, Besançon, France; version ST 8.1.9286) to process the data acquired from the quantitative use-wear analysis. Here, the template was applied to a *Keilmesser* from Balver Höhle (MU-232). This is page 5 of 5.

parameter	unit	category	subcategory	name	description
Smc (mr = 10 %)	µm	Field	Functional	inverse areal material ratio of the scale-limited surface	height *c* at which a given areal material ratio (*mr*, or *p* in MountainsMap) is satisfied
Smr (c = 1 µm under the highest peak)	%	Field	Functional	areal material ratio of the scale-limited surface	ratio of the area of the material at a specified height *c* to the evaluation area
Sxp (p = 50 %, q = 97.5 %)	µm	Field	Functional	peak extreme height	difference in height between the p and q material ratios
Sa	µm	Field	Height	arithmetical mean height of the scale limited surface	arithmetic mean of the absolute of the ordinate values within a definition area (*A*)
Sku	<no unit>	Field	Height	kurtosis of the scale-limited surface	quotient of the mean quartic value of the ordinate values and the fourth power of Sq within a definition area (*A*)
Sp	µm	Field	Height	maximum peak height of the scale limited surface	largest peak height value within a definition area
Sq	µm	Field	Height	root mean square height of the scale-limited surface	root mean square value of the ordinate values within a definition area (*A*)
Ssk	<no unit>	Field	Height	skewness of the scale-limited surface	quotient of the mean cube value of the ordinate values and the cube of Sq within a definition area (*A*)
Sv	µm	Field	Height	maximum pit height of the scale limited surface	minus the smallest pit height value within a definition area
Sz	µm	Field	Height	maximum height of the scale-limited surface	sum of the maximum peak height value and the maximum pit height value within a definition area
Sdq	<no unit>	Field	Hybrid	root mean square gradient of the scale-limited surface	root mean square of the surface gradient within the definition area (*A*) of a scale-limited surface
Sdr	%	Field	Hybrid	developed interfacial area ratio of the scale-limited surface	ratio of the increment of the interfacial area of the scale-limited surface within the definition area (*A*) over the definition area
Sal (s = 0.2)	µm	Field	Spatial	autocorrelation length	horizontal distance of the $f_{ACF}(t_x,t_y)$ which has the fastest decay to a specified value *s*, with $0 \leq s < 1$
Std (Reference angle = 0°)	°	Field	Spatial	texture direction of the scale-limited surface	angle, with respect to a specified direction θ, of the absolute maximum value of the angular spectrum
Str (s = 0.2)	<no unit>	Field	Spatial	texture aspect ratio	ratio of the horizontal distance of the $f_{ACF}(t_x,t_y)$ which has the fastest decay to a specified value *s* to the horizontal distance of the $f_{ACF}(t_x,t_y)$ which has the slowest decay to *s*, with $0 \leq s < 1$
Vm *(p = 10 %)*	µm³/ µm²	Field	Volume	material volume	volume of the material per unit area at a given material ratio calculated from the areal material ratio curve
Vmc (p = 10 %, q = 80 %)	µm³/ µm²	Field	Volume	core material volume of the scale-limited surface	difference in material volume between the *p* and *q* material ratios

parameter	unit	category	subcategory	name	description
Vmp (p = 10 %)	µm³/ µm²	Field	Volume	peak material volume of the scale-limited surface	material volume at *p* material ratio
Vv (p = 10 %)	µm³/ µm²	Field	Volume	void volume	volume of the voids per unit area at a given material ratio calculated from the areal material ratio curve
Vvc (p = 10 %, q = 80 %)	µm³/ µm²	Field	Volume	core void volume of the scale-limited surface	difference in void volume between *p* and *q* material ratios
Vvv (p = 80 %)	µm³/ µm²	Field	Volume	dale void volume of the scale-limited surface	dale volume at *p* material ratio
Asfc	<no unit>	SSFA		Area-scale fractal complexity	The slope of the steepest part of the curve, fit to a log-log plot of relative area over the range of scales multiplied by -1000
epLsar	<no unit>	SSFA		Exact proportion Length-scale anisotropy of relief	Radial profiles are extracted from the centre of the surface, by default every 5°, and the value of relative-length is calculated at a scale of 1.8 µm
HAsfc	<no unit>	SSFA		Heterogeneity of Area-scale fractal complexity	Calculated in a scale-sensitive manner by splitting individual scanned areas into successively smaller subregions given equal numbers of rows and columns
NewEpLsar	<no unit>	SSFA		Exact proportion Length-scale anisotropy of relief	Radial profiles are extracted from the centre of the surface, by default every 5°, and the value of relative-length is calculated at a scale of 1.8 µm
Smfc	µm²	SSFA		Scale of maximum complexity	The scale range over which *Asfc* is calculated (the steepest part of the relative area versus scale curve)
Isotropy	%	Texture direction		Isotropy	Identical to ISO 25178 *Str*
Tr1R	°	Texture direction		First direction	Only relevant if Isotropy ≤ 30 %. Identical to ISO 25178 *Std*
Tr2R	°	Texture direction		Second direction	Only relevant if *Isotropy* ≤ 30 %
Tr3R	°	Texture direction		Third direction	Only relevant if *Isotropy* ≤ 30 %
madf	µm	Furrow		Maximum depth of furrows	Furrows are micro-valleys. The furrow analysis identifies the areas where points are lower than the neighbouring points on a given surface
metf	µm	Furrow		Mean depth of furrows	Furrows are micro-valleys. The furrow analysis identifies the areas where points are lower than the neighbouring points on a given surface
medf	cm/cm²	Furrow		Mean density of furrows	Furrows are micro-valleys. The furrow analysis identifies the areas where points are lower than the neighbouring points on a given surface

Table 11 3D parameters calculated within the qualitative use-wear analysis. These 3D parameters include 21 ISO 25178-2 parameters, 3 furrow parameters, 3 texture direction parameters, 1 texture isotropy parameter and 6 SSFA parameters.

CONTROLLED EXPERIMENTS

Assigning a function to Palaeolithic artefacts requires profound technological and morphological knowledge as well as suitable analogies and comparisons. While ethnographic comparisons are most likely revealing, they do not provide a one-to-one interpretation (Binford 1968). Determining a tool's function has proven to be especially complicated in time periods, when toolkits were not highly specialised and tools used for multifunctional purposes as for example in the Middle Palaeolithic (Plisson et al. 1998). In order to address tool function, experiments provide an alternative possibility to test assumptions and gather new information. Experiments have played an important role in archaeology since the 1970s (Tringham et al. 1974; Coles 1979; Odell/Odell-Vereecken 1980; Outram 2008). The general idea behind the conduction of experiments is to reconstruct past processes and to replicate potential past human activities. Experiments also frequently serve the purpose to produce reference collections for comparison with archaeological samples. Although different experiments follow different research questions, they all have something in common: they all test hypothesis. Hypothesis testing is used to assess the plausibility of a hypothesis (null hypothesis) by using sample data. A test can thereby only provide evidence concerning the plausibility of the hypothesis and the conclusion cannot be proven as true. A tested assumption can only be accepted as temporarily valid.
Depending on the question addressed by an experiment, different levels or types of experiments can be conducted. The distinction between »*Actualistic*«, »*Pilot*« or »*Exploratory*« on the one hand and »*Controlled*«, »*Second generation*« or »*Laboratory*« experiments on the other hand can be made (Eren et al. 2016; Marreiros/Pereira/Iovita 2020). The first type of experiments (»*Actualistic*«*,* »*Pilot*« or »*Exploratory*«) aims mainly at replicating human actions in realistic contexts, whereas the second type focuses on testing individual variables *(*»*Controlled*«*,* »*Second generation*« or »*Laboratory*«). These experiments often involve mechanical devices in order to limit the human bias. By isolating certain variables, their effect on the final result can be comprehended. Recently, an attempt has been made to refine this terminology (Marreiros et al. 2020). Based on this, a distinction of three generations of experiments is reasonable.

First generation experiments

First generation experiments are comparable with the prior described »*Actualistic*«, »*Pilot*« or »*Exploratory*« experiments. These experiments build the foundation for new hypothesis. *First generation experiments* often test the possibility to perform a certain action with a tool. With this, they provide information in order to help understanding past technologies. *First generation experiments* are a crucial step for making preliminary observations and the identification of major variables. Nevertheless, these experiments cannot go further than delivering insights into certain aspects concerning for instance, tool performance.

Second generation experiments

Second generation experiments do not seek to answer overarching research questions, instead they are focused on basic fundamental aspects. This often concerns uniform principles for instance physical principles, operating uniformly across space and time. The identification of patterns, processes and the test of key properties should be the focus of the experiments. Therefore, the archaeological interpretations should be detached from the data interpretation. Performing a *second generation experiment* means testing individual variables. With this approach, the effect of the individual variables in the setup can be evaluated

respectively. Furthermore, the aim of *second generation experiments* is to control and manipulate actions. The design of these experiments is thus more complex and involves mechanical devices. Human variability should be reduced to a minimum. By following this approach, not only the performed action is standardised through a mechanical apparatus, but also the samples and the contact material. Ideally, the samples imitating archaeological artefacts or at least certain features are standardised in order to guarantee variable control and comparability. The same applies to the contact material. *Second generation experiments* reduce subjectivity and result in information about the cause and effect of the tested variables, but at the same time, they are further away from replicating human actions and thus the results cannot directly be transferred to the archaeological record. Since *second generation experiments* are controlled and standardised, they also seek to produce repeatable and reproducible results.

Third generation experiments

Experiments from the *third generation* build on the results gained through *second generation experiments*. In this step, the human variability should again be incorporated. The recognised patterns and the generated models from the *second generation experiments* should be tested in a more naturalistic way. Mechanical devices are not needed in this type of experiment. Instead, the use of so-called multi-sensor systems for gesture recognition (Pfleging et al. 2015; Key 2016; Williams-Hatala et al. 2018) is recommended. Standard samples and replicas of archaeological tools should be employed in an identical way. This approach makes it possible to correlate the results. The *third generation experiments* allow evaluating the results from the *second generation experiment* by adding the human variability and bias. Doing so, the results can have implications on the interpretation of archaeological assemblages and can address overriding research questions. Archaeologists have conducted experiments for decades. Since then, criticism has also accumulated. In particular *controlled* or *second generation experiments* are exposed to scepticism. As a major point of criticism, the artificial nature of these experiments has to be highlighted. The division in the three generations of experiments is trying to overcome this problem. It needs to be stressed that the different generations do not exclude each other, on the contrary; they should be seen as complementary. To go further, for designing a *second generation experiment*, several researchers require keeping the following conditions: the research question should be defined clearly, including a hypothesis as well as an alternative hypothesis which can be tested. The number of trials and samples has to be high enough in order to be statistically evaluable. This should also lead to a dominant use of quantitative over qualitative methods. The involved variables need to be identified and assigned as independent, dependent and random variables. The experimental setup has to be detailed in the explanations in corresponding publications. Together, *second generation experiments* should target at repeatability and reproducibility.

Although discussions concerning experiments, especially *second generation experiments* are becoming numerous and first improvements are visible, there are limitations. In general, sequential experiments are time consuming. Moreover, mechanical devices are expensive and their incorporation in an experimental design is most commonly unrealistic. Even if the use of a mechanical device is possible, testing, manipulating and controlling each variable related to an archaeological assemblage is extremely elaborated. The use of quantitative methods and open data in archaeological research is unfortunately not yet common practice. Thus, cooperation and data exchange in order to minimise the costs and the time investment are difficult to implement.

The controlled experiments conducted within this project belong to the so-called *second generation experiments*. These experiments aim to address basic questions and to understand the influence of specific

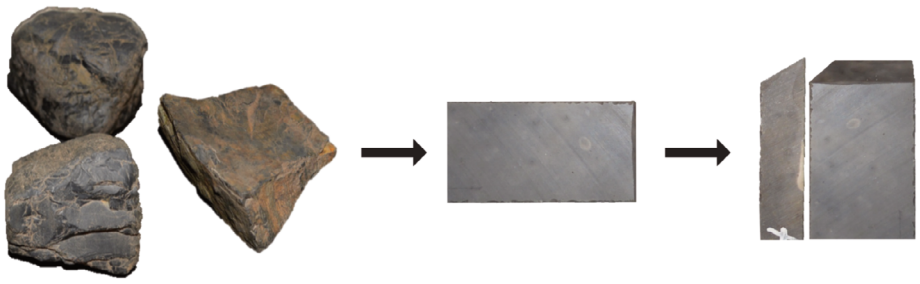


Fig. 35 Standard sample production. The raw pieces (left; here silicified schist) are cut into blanks (middle) with a rock-saw. A diamond band-saw was used to cut the blank in order to create a typical standard deign sample with a defined edge angle (right).

variables within a complex system. Thus, variables need to be tested individually in order to test their effect-causation with other variables in the experiment. In *second generation experiments*, the human factor as well as subjectivity are reduced to a minimum. To do so, a mechanical device was used to perform defined actions. This mechanical device involved in the controlled experiments here is a modular material tester (SMARTTESTER®, Inotec AP GmbH, Germany). The SMARTTESTER® allows for precise movements and is sensor-controlled (Calandra/Gneisinger/Marreiros 2020). Using the mechanical device makes the experiment itself repeatable. To further improve this aspect and to actually isolate certain variables, standard samples were used during the experiments. Raw-material intra-variability was limited by the use of as few as possible raw material nodules. The standard samples are machine cut samples with a predefined shape (**fig. 35**). Baltic flint as well as silicified schist served as raw material. The raw material, in case of the silicified schist, was collected next to the small streams in the surrounding areas of the two sites Buhlen and Balver Höhle. The silicified schist and flint nodules were cut with a rock-saw into cubes, and further cut into size-defined blanks. While the sample length could vary, the standard width (length edge) of the samples was always 3 cm. The cubes were separated into as many blanks as possible to reduce raw material variation. In the next step, one end of the blanks was cut with a diamond band-saw (EXAKT 310 CP) in a given angle. One side of the active edge was in addition modified by a 45° chamfered edge (exception: initial experiment). The chamfered edge was meant to reduce the force applied on only one point of the sample when the contact with the worked material was initialised. The idea was, that this modification would spread the applied force on a bigger area and would therefore minimise the risk of immediate fracture. This cut was also done with the diamond band saw. Usually, the cut with the band saw left a small burr between the two adjacent surfaces – the surface of the active edge and the chamfered edge. This burr was manually removed with a mini diamond drill bit. The samples were cleaned after cutting. First the samples were rinsed with tap water and then with a preheated ultrasonic bath (EMAG Emmi 20HC). For that reason, the samples were individually packed in a plastic bag filled with ~ 100 ml of demineralised water and a non-ionic detergent (BASF Plurfac LF901, 1 g/l = 1 % w/v; BASF SE, Ludwigshafen, Germany). The samples were left in the bath for 10 min at 45°C and 100 KHz. Subsequently, the demineralised water was exchanged with tap water. This was repeated two more times. In a final step, the bags were filled with ~100 ml purified water and emptied after rinsing. Since it was crucial for the later use-wear analysis to enable the analysis of the exact same surface, a coordinate system was created directly on the sample (Calandra et al. 2019a). For this, three 100-200 µm diameter ceramic beads were adhered with epoxy resin (Epoxydharz L mixed with Härter S in a ratio 10:4 by weight) on the dorsal and ventral site of the tool.

The experiments were described with the aim to address the same overall research question. The goal is to bring the relation between tool morphology and tool function into question.

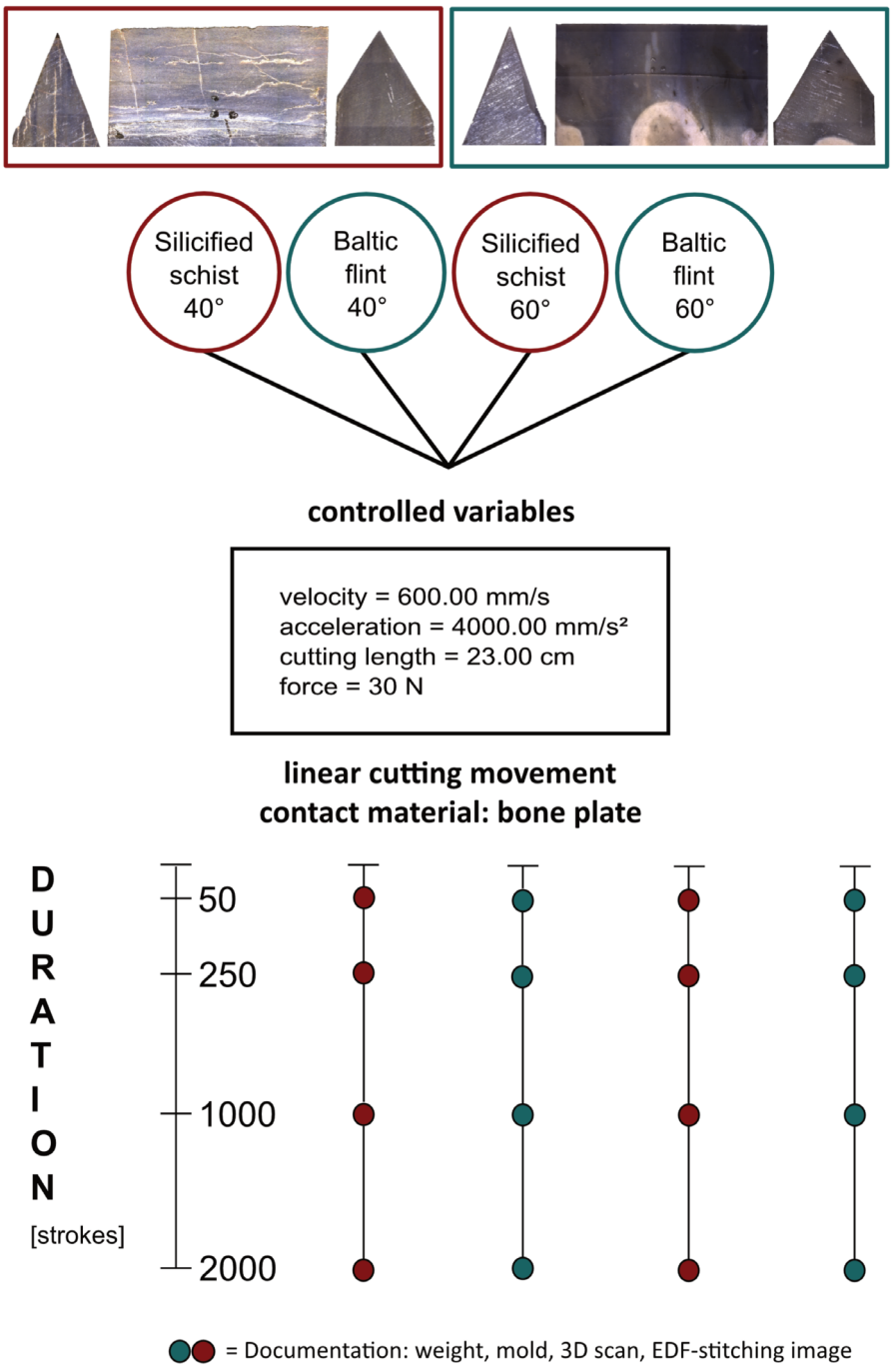


Fig. 36 Experimental design for the initial experiment.

Experimental setup – initial experiment

The initial experiment was designed to evaluate the experimental setup and to test two variables, raw material and edge angle (**fig. 36**). Silicified schist as well as Baltic flint served as raw material. The standard samples were bifacially cut as 40° and 60° degree edge angles respectively. The experiment had no demand on statistical validity. With a sample size of n = 7 (3× silicified schist 40°, 2× silicified schist 60°, 1× flint 40°, 1× flint 60°) not enough replicas were included in order to lead to statistically significant results. The electric, linear drive (z-direction) of the mechanical devise, the SMARTTESTER®, was used in order to perform a unidirectional cutting movement. As a prerequisite in order to standardise the experiment as much

as possible and to test the effects of the two mentioned variables individually, the contact material was also standardised. Therefore, an artificial bone plate (PR0114) from the Swiss company SYNBONE® was used. This generic plate is a modified bone-like polyurethane with the dimensions 250 mm × 250 mm × 6 mm and a shore hardness (D) of 78 5/- 5 %.
The experiment was built up as a sequential experiment. Four cycles defined by a number of cutting strokes were executed. The first cycle started from zero to 50 strokes, the second from 51-250 strokes, the third 251-1000 strokes and the final cycle from 1001-2000 strokes. The cutting length was thereby 23 cm. Each sample had to perform the four cycles with 2000 strokes in total by always using the same cutting track. Certain factors were set up during the experiment, others were measured by sensors only. The velocity with 600.00 mm/s and the acceleration 4000.00 mm/s belong to the predetermined factors. The force of 30 N was also defined by the weight of the sample holder (~ 1 kg) and two manually added kg. In total five sensors were activated during the entire duration of the experiment. Three sensors recorded the predetermined factors velocity, acceleration and force. The other two sensors recorded the penetration depth in mm and the friction in N. The sampling rate was in a frequency of ten Hz. The samples were clamped straight into the sample holder in a 90° angle between sample and contact material. A template per sample was created to assure the constancy of the programmed settings as for instance the position of the sample on the x-, y- and z-axis.
To test aspects such as efficiency, performance and durability by comparing the two raw materials and the two edge angle values, the documentation of the samples needed to follow a protocol. Initially, the samples were photographed with a Nikon D610. The weight was recorded with a weighing scale, Kern PCB 3500.2, with an accuracy of 0.1 g. To avoid inaccuracies, each sample was measured three times. The material properties of the raw material were tested using the Leeb rebound hardness tester (Proceq Equotip 550, Leeb C probe) by means of a coupling paste. The samples were scanned with an AICON smartScan-HE R8 from the manufacturer Hexagon. The system was always turned on one hour in advance in order to warm up all components. The S-150 FOV used has a resolution of 33µm. Identical settings were used for all scans; scans were exported afterwards in an STL-format. Three of the four surfaces per sample (one lateral and the two main surfaces) were documented with the ZEISS Smartzoom 5, a digital microscope. The 1.6× objective was used to create EDF-stitching images. Moulds were taken from the two cutting surfaces covering the edge. This protocol was not only followed before the experiment, but also after each cycle. Frequently, single cycles were video recorded.

Experimental setup – »artificial VS. natural« experiment

In the sequence of experiments was the so-called »artificial VS. natural experiment« the second one (**fig. 37**). This experiment aimed to validate the comparability of artificial with natural contact material. Tool performance and durability as well as the produced use-wear traces served as a measure of criteria. In order to standardise the contact material during an experiment, these artificial materials are a convenient solution, but their use needs to be justified. In addition, the experiment aimed at understanding the development of use-wear traces and testing the mechanical trends behind the formations.
Within this experiment, the independent variables were the raw material – silicified schist and Baltic flint – and the contact material. Four different contact materials were tested. Two belonging to the category artificial material and the other two were natural, fresh materials. An artificial bone plate coated with a rubber skin (PR0114.G) from the Swiss company SYNBONE® was used. This generic plate is a modified bone-like polyurethane with the dimensions 250 mm × 250 mm × 6 mm and a shore hardness (D) of 78 5/- 5 %. The

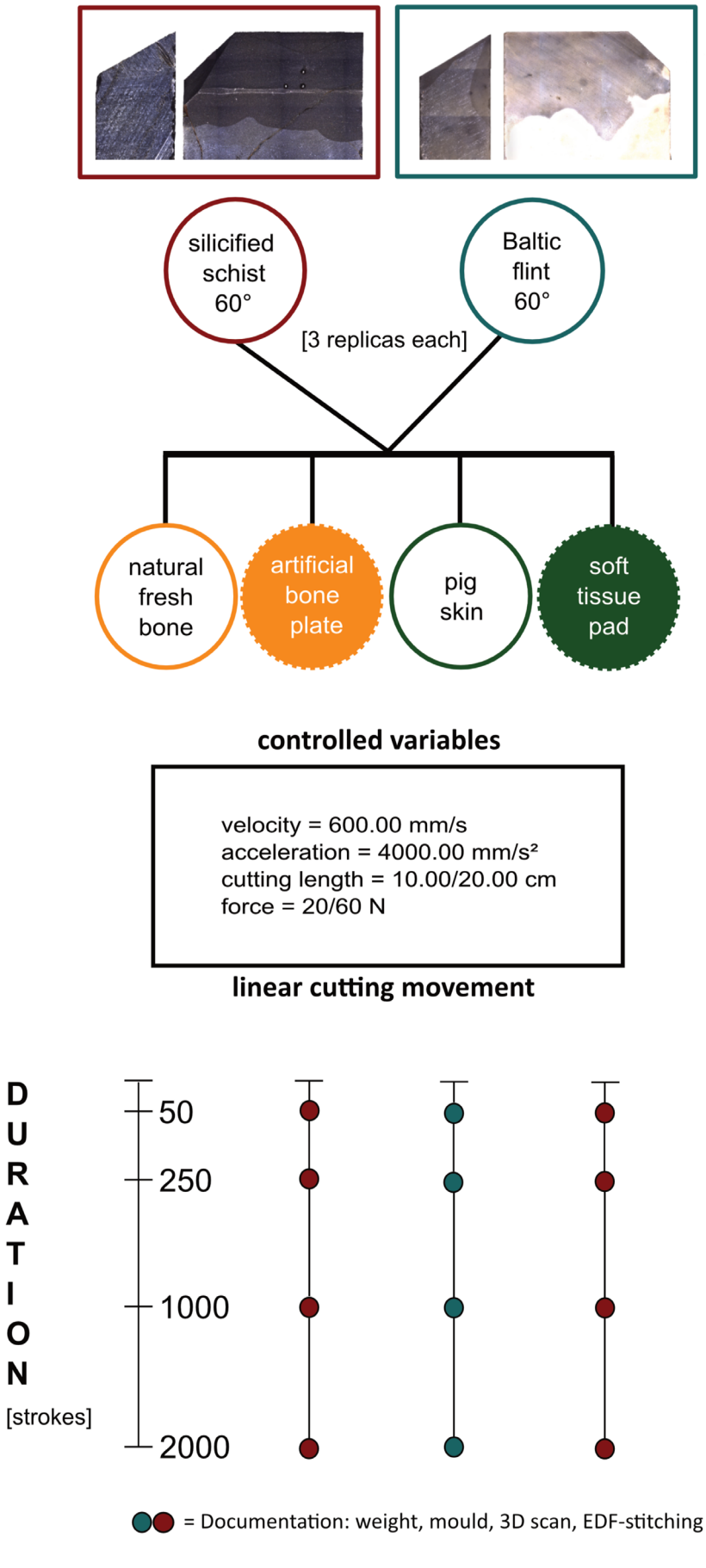


Fig. 37 Experimental design for the »artificial VS. natural« experiment.

bone plate with the rubber skin layer, imitating the periosteum, was therefore an artificial equivalent to a fresh and defleshed bone. The bone in this experiment was a cow scapula (*Bos primigenius Tarurus, Angus*) provided in a fresh state by a butcher. The periosteum and small pieces of flesh were still attached. The experiment was carried out in a laboratory under ambient room temperature conditions. During the course of the experiment, the bone started to dry out but apart from desiccation no other degradation were visible. The reason to choose a cow scapula can be explained by the size and shape. The morphology of the bone offers the possibility to perform long cutting strokes on a considerably straight surface, providing better conditions for a comparison of the results between the bone plate and the bone. The second set of contact material was skin. As artificial skin a soft tissue pad with a matrix (PR1043.10) SYNBONE® was used. This skin pad is made of silicone ecoflex and has the dimensions 140 mm × 130 mm × 4.2 mm and a shore hardness 0-3A. A piece of natural pork skin was provided by a butcher by separating the skin from the flesh below. The skin was kept outside under room temperature and overnight it was kept cold in a fridge. Two pieces of skin were needed during the course of the experiment. The n = 24 samples were cut unifacailly with an edge angle of 60° and equally distributed between the four contact materials. Thus always three replicas per raw material were tested on one contact material.
The setup for this experiment was identical with the setup from the initial experiment. The same linear drive was used in order to perform unidirectional cutting strokes, the same number of strokes separated in four cycles was executed and the same sensors with identical settings were used. A few changes however, are noteworthy. For the bone plate and the natural bone a force of 60 N was applied, while the force for the artificial and natural skin was only 20 N. The cuts on the bone plate had a length of 20 cm and half of the cuts on the fresh bone, too. The other cutting strokes were ten cm long, but the amount of strokes was doubled in these cases. The documentation followed the same protocol mentioned for the initial experiment.

Experimental setup – tool function experiment

The third and main experiment built up on the results carried out during the other two experimental setups (**fig. 38**). The goal was again to test efficiency, durability and performance by comparing two raw materials and different edge angles on standardised contact material. To further investigate the relation between tool morphology and function, the edge angle values for the standard samples were extrapolated from the edge angle measurements (see chapter Quantification of edge design) from the sampled *Keilmesser*. These measurements are based on 3D scans of in total 157 *Keilmesser*. The selected tools were separated in *Keilmesser* (n = 57) and *Keilmesser* with a modification through the application of the *Prądnik method* (n = 100). From all the calculated edge angle data only the results from the »2-lines« procedure with measurements at 5 mm distance to the intersection were selected. The statistical analysis was performed with the open-source software R version 4.0.2 through RStudio version 1.3.1073 (RStudio Inc., Boston, USA) for Microsoft Windows 10. The following packages were used: writexl v. 1.3, tidyverse v. 1.3.0, openxlsx v. 4.1.5. and R.utils v. 2.9.2. Reports of the analysis in HTLM format, created with knitr v. 1.29 and rmarkdown v. 2.3 are available on GitHub [https://github.com/lschunk/edge_angle_analysis] as well as the raw data, the scripts and the RStudio project. For calculating the average edge angle on the selected tools, the first and the last horizontal section (always ten sections per tool) were excluded (**fig. 39**). This is reasoned by the fact that these sections usually display edge angle values distinct from the values of the other eight sections. These two sections are likely not representing the active edge anymore. It is possible that they are already part of the base and the arch of the tool. Based on the eight horizontal sections and the measurements at 5 mm the average per tool could be calculated. Since there was a notable change in the edge angle values in the distal

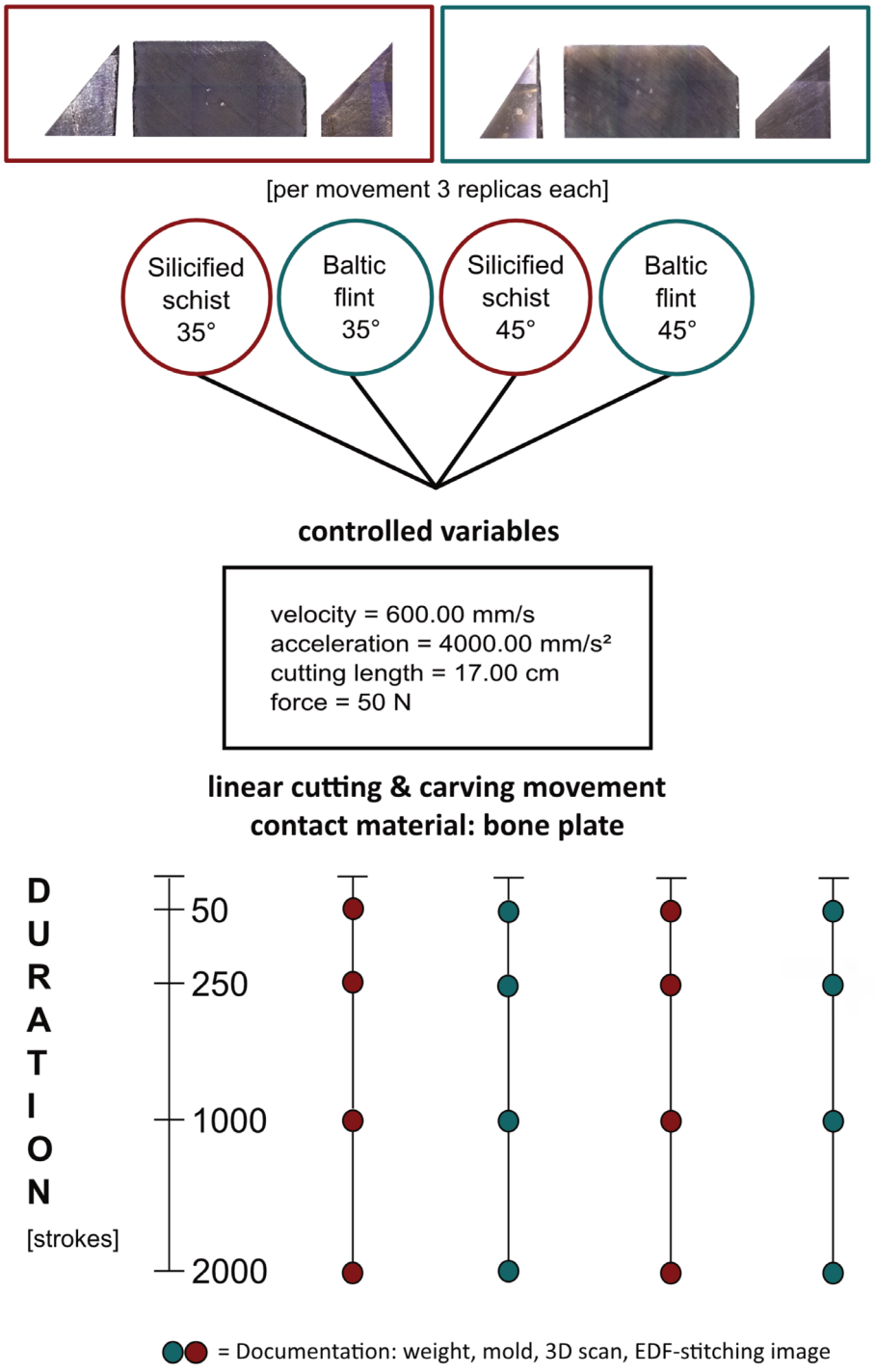


Fig. 38 Experimental design for the tool function experiment.

and the proximal part of the tool, two instead of one average value were calculated. The first built on the results of the section two to four and the second on the results of the section five to nine. The average was calculated per classification (*Keilmesser* and *Keilmesser* with *Prądnik method*). The results for the proximal tool part were nearly identical in the two classifications (44.75° and 44.48° respectively). For the distal part of the tool the results varied slightly. While *Keilmesser* reach an average of 36.06°, *Keilmesser* modified by the *Prądnik method* show an average of about 33.50°. Therefore, the decision was made to test two different edge angle values – 45° and 35° – as a proxy for the proximal and the distal part of the tool, respectively.

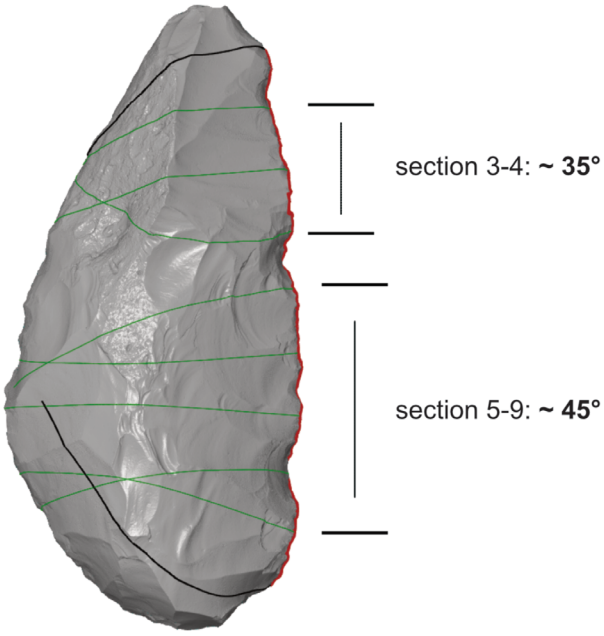


Fig. 39 Calculation of the edge angles, illustrated here on a 3D model of a *Keilmesser* from Ramioul (ID R-019). The horizontal lines are the section. The sections used to calculate the edge angels in green, in black the ones excluded for the calculation. The edge angle values have been calculated with the »2-lines«.

In total 24 standardised samples were unifacially cut. Twelve of these 24 samples were produced out of Baltic flint nodules, the other twelve from silicified schist blocks. The linear drive was used to perform two different movements: unidirectional cutting and carving. Artificial bone plates served as contact material (PR0114), the same modified bone-like polyurethane plates as used during the initial experiment. The cutting and carving length was 17 cm each. Three samples per raw material and edge angle were tested for cutting and carving respectively. The experimental setup was identical to the ones during the first and second experiment. Thus, the tool function experiment was also designed as a sequential experiment with the same four cycles and 2000 strokes in total per sample. A force of 50 N was applied for both movements. While the samples for the cutting movement were clamped again straight into the sample holder in a 90° angle between sample and contact material, the sample was adjusted differently for the carving movement. With the flat, not angled tool surface, the samples were clamped into the sample holder in a 20° angle towards the contact material. Consequently, the carving movement was performed in a flat and unidirectional motion. The documentation followed the same protocol as mentioned for the initial and the »artificial VS. natural« experiment.

Analysis of the experiments

Since the three experiments were following an identical experimental setup and protocol for the documentation, the acquired date could be evaluated similarly. Furthermore, the data could be combined. To answer questions related to tool efficiency, durability and performance, the samples were review in terms of their morphological changes. Material loss, fractures, retouch and changes in the edge angle were documented. The contact material was evaluated on the one hand concerning material loss, on the other hand the penetration depth was a valid indicator. Use-wear traces resulting from different actions and contact materials on the two different raw materials can be compared as well. Due to the sequential aspect of the experiments, also statements regarding the temporal formation and the development of the use-wear traces can be addressed.

Analysis 3D data

Part of the documentation before and after each cycle of the experiments was the 3D scanning of the samples. In order to further analyse the 3D models, a few preparatory steps were necessary. After scanning, the resulting STL file was imported in GOM Inspect, an open source software for 3D measurement data (GOM

Software 2018, Hotfix 1, Rev. 111729, build 2018-08-22). With GOM inspect, the 3D models were checked for holes in the mesh. Existing holes were closed automatically with a maximum hole size of 10.000 mm and 1000 as maximum number of edges. Mesh errors were eliminated. Depending on the number or the size of the holes, these two steps were repeatedly applied. The holes were not closed directly in one step with a higher maximum hole size value in order to avoid a change in the original morphology of the scanned sample.
Based on the closed models, the volume per sample could be computed. In this way, material loss of the samples due to the use within the experiment can be detected. At the same time, the 3D models can be inspected visually. GOM inspect allows for a comparison of 3D meshes. For this step, one of the meshes needs to be defined as CAD body (computer-aided design), while the other one can be imported as a mesh. The two models can be optically combined with a local best-fit alignment. Afterwards, a surface comparison on the CAD can be initialised. Based on a colour scale, the differences between the two 3D meshes can be optically seen and also numerically read.

Analysis edge angle

The closed 3D models also serve for the edge angle calculation (see method section 4.3). A surface curve defining the edge of each sample was added in GOM inspect. This polyline was exported in an IGES-format. All further steps were identical as described in the section »Quantification of edge angle design« earlier in this chapter. The following parameters were chosen: The number of horizontal sections is represented by ten sections per tool. The measurements were taken in one mm steps starting with the intersection between the vertical and the horizontal polylines. The lengths of the lines used in the »2-lines« and »best fit« measuring procedure, was always 2 mm. Calculating the edge angle of the experimental samples allows for a quantification of a change in the edge angle.
The statistical analysis was performed with R (version 4.0.2 through RStudio version 1.3.1073, RStudio Inc., Boston, USA). The following packages were used: writexl v. 1.3, tidyverse v. 1.3.0, openxlsx v. 4.1.5, ggrepel v. 2.3, doBy v. 4.6.7, patchwork, v. 1.0.1, ggplot2 v. 3.3.2 and R.utils v. 2.9.2. Reports of the analysis in HTLM format, created with knitr v. 1.29 and rmarkdown v. 2.3 are available on GitHub as well as the raw data, the scripts and the RStudio project [https://github.com/lschunk/edge_angle_experiments].

Use-wear analysis

Use-wear analysis was not done for all experimental samples. Instead, the pieces were systematically sampled. None of the samples from the initial experiment were part of the use-wear analysis. In the case of the second experiment, the »artificial VS. natural' experiment, eight samples were selected. These samples consist of four silicified schist and four flint samples. Each of the four samples was tested on a different contact material: bone plate, scapula, skin pad and pork skin. A further eight samples belonging to the tool function experiment were analysed. The selection was based on the raw material of the sample, the edge angle and the movement. Thus, one tool per cutting with 35° and 45° edge angle and one tool per carving with 35° and 45° edge angle made of silicified schist and flint were sampled.
As for the archaeological material, a scheme was designed for the quantitative use-wear analysis as well (**fig. 40**). The scheme fulfilled the purpose of mapping and locating the use-wear traces consistently in an identical way. The outline of the standard samples was used as template for a grid of four areas. This was

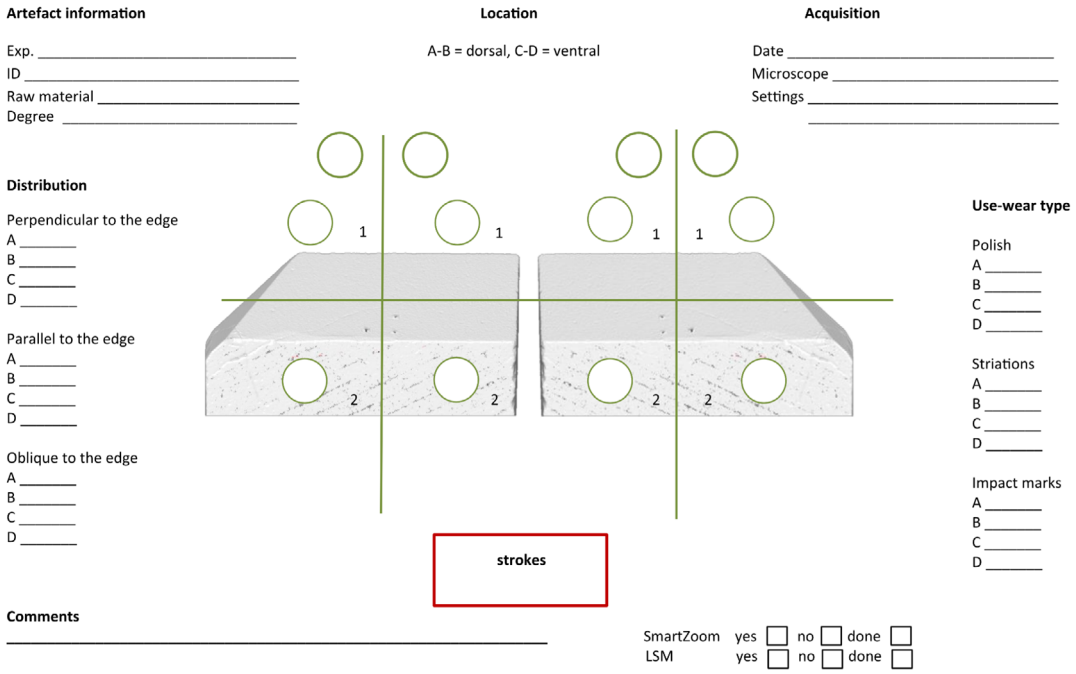


Fig. 40 Scheme used during the qualitative use-wear analysis for the documentation of the analysed standard samples from the experiments. The dorsal surface is indicated by A and B (surface with the angle) and the ventral surfaces by D and C (straight surface). Additionally, the scheme includes information about the analysed standard sample itself, the cycle within the experiment, the acquisition date and settings as well as information about the use-wear traces (location, distribution and type). The scheme also contains information about, whether further acquisition or analysis are needed (SmartZoom = digital microscope images, LSM = quantitative use-wear analysis).

done for the dorsal (A and B) as well as for the ventral (D and C) surface. The areas are numbered from one, as the part of the tool with the edge, to two, the lower part of the tool.

The qualitative use-wear analysis was done in a high-power approach by means of an upright light microscope (ZEISS Axio Scope.A1 MAT). The samples were studied with a 5×, 10× and 20× magnification. Traces were documented as an EDF image in black and white settings. While the eight samples from the »artificial VS. natural« experiment were analysed before and after each cycle, the eight samples from the tool function experiment were only analysed before and after the last cycle with 2000 strokes.

The quantitative use-wear analysis was done on the same total 16 samples, using the laser-scanning confocal microscope (ZEISS Axio Imager.Z2 Vario + ZEISS LSM 800 MAT). The objective C Epiplan-Apochromat 50× with the numerical apertures 0.95 was generally used. In a few cases, the C Epiplan 50× /0.75 had to be used. The field of view (FOV) was 255.6 × 255.6 µm. For the quantitative analysis, measurements were taken on samples only from before and after 2000 strokes, in the case of the tool function experiment only after 2000 strokes. On each sample, one use-wear spot was measured three times at a nearby but non-identical point. Only on two samples, a second use-wear spot was measured as well. In total, 28 measurements were taken on the experimental sample times three. Since the qualitative as well as the quantitative use-wear analysis was done after the experiments were finished, the measurements were taken from the moulds. Exceptions are the samples after 2000 strokes from the tool function experiment. In these cases, the original surface not the replicated mould's surface were measured.

The reason for doing the use-wear analysis on the experimental samples was firstly to see what kind of use-wear develops within the sequential and automated experiments and also to see how the use-wear created differs depending on the contact material used. The second important aspect was to observe how use-wear develops in general and also under the application of the two tested raw materials. Therefore,

each time it was important to measure the exact same spot before and after the experiments. This was possible due to the coordinate system on the samples. The use of the coordinate system should guarantee that a possible measured variability in the surfaces topography is only due to a surface alteration and not because of a positional inaccuracy. The LSM has an integrated module called Shuttle-and-Find (Calandra et al. 2019a). This module allows the relocating of a sample based on a defined coordinate system. Prior to a measurement, the coordinate system needed to be created and calibrated for each sample with the C Epiplan-Apochromat 10× /0.40. Firstly, the surface of the sample after 2000 strokes was analysed. This way, the spot with the use-wear trace could be defined and measured. Afterwards, the sample was exchanged with the corresponding sample from before the experiment. Based on the previously calibrated coordinate system, the software could automatically relocate the same spot on the surface with a horizontal positional repeatability of approximately 14 % of the FOV.
The acquired data was processed identically to the data form the quantitative use-wear analysis of the archaeological material (see chapter Quantitative use-wear analysis). The data of the two experiments was individually processed with templates in ConfoMap (a derivative of MountainsMap Imaging Topography developed by Digital Surf, Besançon, France; version ST 8.1.9286). The maximum of two templates had to be applied. The first template, the »resampling template«, was used to perform a resampling in x and y on all 108 surfaces, leading to an identical spacing. In the second and final template, that data was processed (»processing template«). The following procedure was performed: I.) Levelling (least squares method by subtraction), II.) From removal (polynomial of degree3), III.) Outliers removal (maximum slope of 80°), IV.) Thresholding the surface between 0.1 and 99.9 % material ratio to remove the aberrant positive and negative spikes, V.) Applying a robust Gaussian low-pass S-filter (S_1 nesting index = 0.425 µm, corresponding to about five pixels, end effects managed) to remove noise, VI.) Filling-in the non-measured points (NMP), VII.) Analysis: Calculation of 21 ISO 25178-2 parameters, three furrow parameters, three texture direction parameters, one texture isotropy parameter and the SSFA parameters *epLsar, Asfc, Smfc, HAsfc9 and HAsfc81*. The ConfoMap templates for both experiments in MNT and PDF formats are available on GitHub, including all original and processed surfaces, as well as the results.
Again, the open-source software R (version 4.0.2 through RStudio version 1.3.1073, RStudio Inc., Boston, USA) was used to perform all descriptive analyses. This includes the summary statistics, the scatter plots and PCA's. As packages openxlsx v. 4.1.5, R.utils v. 2.9.2, doBy v. 4.6.7, ggsci v. 2.9, patchwork v. 1.0.1, ggplot2 v. 3.3.2, tidyverse v. 1.3.0, wesanderson v. 0.3.6 were used. Reports of the analysis in HTLM format, created with knitr v. 1.29 and rmarkdown v. 2.3 are available on GitHub [https://github.com/lschunk/artificial_VS_natural-experiment; https://github.com/lschunk/tool_function-experiment]. Also the raw data, the scripts and the RStudio project are saved in the same repository.

Analysis sensor data

The sensor data recorded throughout the experiments with the SMARTTESTER® can be analysed. Three of the five sensors were predetermined factors: velocity, acceleration and force. This means, the sensors have to be seen as a controlling mechanism. The analysis of these sensor data shows how controlled the experiments were in the sense of permanently repeating actions. The other two sensors recorded the penetration depth in mm and the friction in N. While the friction is a topic for itself and goes into the direction of physics and mechanics and was therefore excluded from further analysis, the penetration depth was analysed. The penetration depth is of interest because it directly relates to aspects of tool efficiency and performance. During the experiment, the sensor data is saved per strokes and per sensor individually as TXT files.

By means of R (version 4.0.2 through RStudio version 1.3.1073, RStudio Inc., Boston, USA) the data from the »artificial VS. natural« experiment as well as the data from the tool function experiment was combined in one data frame respectively and exported as a XLSX file. The data was then statistically analysed and plotted. The following packages were used openxlsx v. 4.1.5, R.utils v. 2.9.2, doBy v. 4.6.7, ggrepel v. 0.8.2, patchwork v. 1.0.1, ggplot2 v. 3.3.2, tidyverse v. 1.3.0 were used. Reports of the analysis created with knitr v. 1.29 and rmarkdown v. 2.3 in HTLM format are available on GitHub [https://github.com/lschunk/Smarttester-data.git]. In the same repository the data, the scripts and the RStudio project can be found.