ELSEVIER

Journal of Archaeological Science 33 (2006) 987–998

Journal of Archaeological SCIENCE

http://www.elsevier.com/locate/jas

Expediting and standardizing stone artifact refitting using a computerized suitability model

Judith R. Cooper [a,*], Fang Qiu [b,1]

[a] Department of Anthropology, Southern Methodist University, Box 750336, Dallas, TX 75275-0336, United States
[b] Program in Geographic Information Sciences, University of Texas at Dallas, TX 75083-0688, United States

Received 10 August 2005; received in revised form 8 November 2005; accepted 11 November 2005

Abstract

Stone artifact refitting is a valuable aspect of archaeological research and can inform on a variety of issues, such as prehistoric technology, site taphonomy, and assemblage patterning and function. It offers a means of teasing apart sites with complicated occupational histories and is particularly useful in interpreting surface lithic scatters, the dominant site type across much of the globe. Unfortunately, refitting is also labor intensive and time-consuming, especially for the inexperienced refitting analyst, making it logistically challenging in the case of many research projects. A possible solution is proposed by which the process of refitting might be partially automated. A multivariate suitability model was created in a Geographic Information Systems (GIS) environment. The refitting suitability model first eliminates low probability refits, and then ranks the remaining artifacts according to a score that reflects their likelihood of refitting to a target artifact. Scores are assigned to assemblage items based on a series of criteria, including raw material, cortex, size, and spatial proximity. In this pilot study, known refits from 5GN149, a surface lithic scatter in Colorado, USA, were used to test the accuracy of the model. The refitting suitability model correctly placed the known refit at the top of the list of potential refits (i.e., assigned a rank to the known refit ranging from 1 to 10) approximately 32% of the time. This is more refit identifications than would be expected through a process of pair-wise comparisons. Preliminary results suggest that the model has the potential to standardize and expedite the process of refitting.

Keywords: Refitting; Geographic Information Systems; Lithic scatters; Suitability models; Analytic hierarchy process

1. Introduction

Surface lithic scatters dominate the archeological record [29,30]. Throughout prehistory these sites were created and often repeatedly exploited, producing an archaeological deposit of wide temporal breadth. What is left on the surface is a two-dimensional scatter of stone tools and debitage, which could signify either an isolated episode from the past or a palimpsest of many [56]. Surface lithic scatters lack the interpretive qualities so heavily relied on by archaeologists. They are not buried, often lack radiometrically datable materials, are particularly susceptible to post-depositional disturbance, and are vulnerable to looting and attendant loss of diagnostic material. For these reasons, lithic scatters are hard to decipher and frequently overlooked by archaeologists [6,30,57]. And yet, the ubiquitous lithic scatter makes up such a large portion of the archaeological record that it is crucial to understanding prehistoric cultural behavior [29]. While many of the analytical methods used by archaeologists cannot be applied to such sites, stone artifact refitting can. Refitting has the potential to parse complicated site histories and extract meaning from of a seemingly insignificant scatter of flakes.

Refitting has been applied to a range of archaeological problems [7]. The technique involves the refitting of flakes and other by-products of tool manufacture, and has been likened to assembling a three-dimensional puzzle in which some pieces are absent [7,11,25]. Stone artifact refitting has a long history

* Corresponding author. Tel.: +1 214 704 6280.
E-mail addresses: jcooper@smu.edu (J.R. Cooper), ffqiu@utdallas.edu (F. Qiu).
[1] Tel.: +1 972 883 4134.

0305-4403/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jas.2005.11.005

in archaeology, with its earliest applications reported in Europe in the late 19th century [55,58]. Refitting was not regularly used by archaeologists until after the 1960s, when it became common in Old World research [14,16,17]. It is used less frequently in North American archaeology [7,14,35,53].

While the analytical utility of refitting is rarely challenged (but see [34]), the refitting process is time-consuming and, consequently, expensive. It might seem impractical within the framework of North American cultural resource management, where both time and funding are often limited [7]. Therefore, while refitting has the potential to significantly enhance our interpretations of the archaeological record, it is often under-utilized due to practical constraints. If refitting were less time-consuming and labor intensive, then it might be more attractive to archaeologists.

We propose a refitting suitability model by which the process of stone artifact refitting is partially automated. The multivariate model is implemented in a Geographic Information Systems (GIS) environment and systematically and rapidly screens all potential refits to a target artifact. Artifacts are evaluated based on the following attributes: cortex, dorsal scar count, size, condition, portion, spatial proximity, and raw material. Unsuitable refits are eliminated, and those remaining are ranked in accordance with their suitability score. In this pilot study, the refitting suitability model is tested for accuracy on an archaeological control sample of refitted artifacts. The goal of the model is to standardize and expedite refitting so the method can be applied more frequently to archaeological assemblages.

2. Background

2.1. Refitting applications

Past applications of refitting have demonstrated its value and versatility within archaeological research. Refitting can be used to detect the natural and cultural processes that contribute to the formation of an archaeological deposit. Through refitting, the extent of artifact movement caused by post-depositional processes, such as bioturbation [15] or plow disturbance [46], can be quantified [10,28,34,41]. Refitting can also help to identify anthropogenic processes that contribute to patterning within a site. The displacement of artifacts in a refit pair may suggest removal of an artifact from its primary context for a specific use, potentially providing valuable data on behavior [9,27,41].

Refitting has been shown to be a powerful spatial analysis tool as well [59]. Through refitting, stratigraphic divisions identified in buried archaeological contexts can be tested to determine if archaeological layers correspond to real, temporal divisions [4,11,26,57,60]. Careful refitting can also isolate and identify activity areas within a site, such as loci of tool manufacture, resharpening events, or discard events [9,21,25,27]. These analyses can provide information on overall site patterning [7,12,19]. By identifying discrete activity areas, archaeologists can tease-apart occupation histories and assess contemporaneity of components [12]. Refitting even has the potential to link disparate sites within a region, which can shed light on prehistoric mobility and regional landscape use [7,14,50,54].

Just as importantly, refitting can inform on stone tool technology, raw material use, and other aspects of the *chaîne opératoire* [7,9,12,25,38,39,52,62]. Through refitting, the steps performed during tool production and maintenance can be retraced. In fact, if the technology that was used to make tools at a site can be established, it may be possible to determine the age of a site that had previously appeared undatable. Refitting also offers a means of testing archaeologists' ability to correctly classify lithic artifacts [2], since classification is easier once an artifact is refitted into a manufacture sequence. The results of refitting analyses can affect the quantitative and qualitative assessments of a lithic assemblage and thus strengthen overall interpretations [16]. Clearly, refitting can extract valuable information from archaeological assemblages, even when other interpretative clues are absent.

2.2. Problems associated with refitting

Archaeologists often regard the time investment and labor intensity required by refitting to be its greatest shortcoming [9,11,12,34,41]. The number of successful refits achieved during the process of refitting is directly proportional to the amount of time devoted by the analyst [9,34]. An analyst who devotes a considerable amount of time to the process should find more refits than a person who devotes only a brief time. But, the greater the amount of person hours devoted to refitting, the greater the cost, making successful refitting projects expensive (at least in instances of large assemblages). In contract or rescue archaeology, when money and time are often limited, refitting may be impractical [7], or at least difficult to justify [27].

Even when considerable time is devoted to refitting, the success rate, measured as the number of successful refits/total refitting population, may still be low [9,35]. In a survey of refitting projects, the average success rate was approximately 15% [16]. Refitting success will vary according to several factors. First, site type and function affect the nature, completeness, and spatial dispersion of an assemblage, all of which in turn impact refitting success [16]. Sites where flint-knapping took place may have a greater chance of containing refits [60] because they often comprise a high percentage of waste flakes deposited at the origin of manufacture. Early-stage reduction locales also contain larger artifacts with a higher incidence of cortex, both attributes that seem to facilitate finding refits [9,41]. Debitage derived from tool retouch or projectile point manufacture episodes might be more difficult to refit because these artifacts are generally smaller in size and lack distinctive cortical features [27]. Refitting success can also be influenced by the size and completeness of the artifact sample, which is shaped by both site formation processes and also the recovery strategy used at a site [9]. If a greater proportion of a site or assemblage is preserved and/or sampled, it improves the chances of finding successful refits because more pieces of the refitting puzzle are available to the analyst [27,60]. Also, raw material type might influence refitting success because certain materials, such as ones marked by distinct coloration,

inclusions, or banding, may be easier to refit [7,9,11,20,49]. In addition, the amount of time devoted to a refitting project would certainly bear on its success rate. Finally, the ability and past experiences of the refitting analyst might affect the success of a refitting project. It can be expected that a more experienced refitting analyst who is familiar with lithic technology would find a greater number of refits [9].

Individual analysts approach the task of refitting in different ways. For some analysts, searching for refits is a process of trial and error [27], where each item is given a chance to refit another in a series of pair-wise comparisons [9,16]. This approach might be practical for a novice, who has less hands-on experience to guide his/her method, but excessively tedious for a more skilled analyst. Others prefer to subdivide the sample according to unit provenience, raw material, technological stage, or breakage patterns [9,27,31,35,41]. Finally, others approach the problem unsystematically and let visual characteristics alone direct the process. Thus, as if solving a puzzle, an individual follows a strategy most appropriate for him- or herself. In fact, an individual analyst may shift strategies several times during the refitting process.

The refitting process is therefore quite subjective and idiosyncratic. Individual analysts use different information to assist in finding refits, and because the analytical process of identifying refits is conducted largely within the analyst's brain, it is difficult to isolate or document the clues the analyst relied upon. As such, the procedure used by one refitting analyst might differ strikingly from those of another, and thus two analysts attempting to refit the same assemblage might end up with very different sets of refits.

It is problematic that the results of a refitting project are so contingent on the ability and technique of the analyst and that both are rarely explicitly defined because archaeologists use these results to make interpretations about a site. For example, a site with a high refitting success rate is often regarded as undisturbed or thought to demonstrate integrity. Perhaps, though, a high refitting success rate was achieved because a skilled analyst devoted a year of time to refitting the assemblage. If the same assemblage were refit by an inexperienced analyst who devoted only a month to the project, then the results would certainly differ. The resulting refitting rates may only be a consequence of the analyst's skill, the amount of time spent refitting, the strategies employed, or, in part, plain luck [9]. In order to make interpretations about assemblages based on refitting results, better control over the refitting process needs to be established.

Interpretations of refitted assemblages would benefit if the method were standardized and a clear systematic approach were defined. The results of one project could be compared to the results of another, and their differences could be attributed to variation in past behaviors or site taphonomy, not merely to variation in the skill of the refitting analysts.

2.3. Improving the refitting method

Considerable effort has been made in recent decades to establish standardized conventions for presenting and visualizing the results of refitting analyses [10,27,36,59,61]. While these efforts aimed to enrich the interpretations of existing refit data, they did not attempt to improve the efficiency of the refitting process, although the tedium of the process seems to be the greatest deterrent to many archaeologists.

Recently, efforts have been made to facilitate and automate the refitting process using three-dimensional scanning and surface modeling [44,45,51]. Although this technique may someday significantly change stone artifact refitting, at present it is impractical for most archaeologists, who lack the hardware, software, or operational expertise.

We propose an alternate approach in which a computerized program is developed and used to help partially automate the refitting process. In the proposed approach, we utilize analysis functions available through standard GIS software, which is becoming increasingly widespread in archaeological research. Therefore, for most, the proposed model would not require additional expensive equipment or advanced technological competence. The method is intuitive and driven by a straightforward algorithm.

3. Methods

The proposed refitting program uses a suitability model to predict refits to a target artifact. In the refitting suitability model, artifacts within the refitting population undergo a preliminary screening process, where unsuitable refits are eliminated [13]. Remaining artifacts are then ranked according to their likelihood of refitting to the target artifact. The goal is to identify the best refit to the target artifact, and then repeat the process iteratively for successive targets.

The computerized model attempts to digitize the mental decision-making process a person engages in while selecting potential refits for a target artifact [41]. But, unlike the internal mental processes of an individual, which involve an immeasurable set of variables, the program uses a defined, controlled set of variables that can be consistently maintained or manipulated between uses. In addition, the refitting suitability model will produce the same results, regardless of skill or experience of the analyst.

The program was developed using the application developer kit of a popular GIS software package – ArcGIS. Like many site predictive models commonly used in archaeology, the model incorporates multiple layers of spatial and attribute data. By using GIS, we were able to organize, display and analyze both spatial and attribute data types. The model requires the data be manually entered by the analyst and then converted to a shapefile format, from which the program can retrieve necessary information to perform its calculations.

3.1. Refitting criteria

Evaluation is based on a series of variables (Table 1), standard to lithic analysis (in part after Andrefsky [3]), coded during the initial artifact analysis. These include (1) *percent cortex on dorsal surface*, coded in four classes (0%; 1–49%; 50–99%; and 100%); and (2) *dorsal flake scar count*, also

Table 1
Variables used to evaluate potential refits

(1) Percent cortex on dorsal surface
(2) Dorsal flake scar count
(3) Artifact size (length)
(4) Condition (broken or complete)
(5) Portion (proximal, medial, distal, unspecified, complete)
(6) Distance to target artifact
(7) Raw material class

coded in four classes (0; 1; 2; 3 or more). Artifact size is coded and can correspond to any relevant measurement (e.g., length, width, area, or weight). For our discussion, (3) *length* is used and was divided into three classes (0–19 mm; 20–39 mm; 40 mm or more). (4) *Artifact condition* (broken or complete) and (5) *artifact portion* (proximal, distal, medial, unspecified, or complete) are coded as well. An artifact's (6) *spatial location* must also be recorded, ideally through piece-plotting, to produce a coordinate location defined by a northing and easting relative to a single site datum. Unit-provenienced artifacts may be used, but because these data are less specific and the same spatial location may correspond to more than one artifact, they produce less discriminating results. The Euclidean distance between each user-selected target artifact and all other artifacts is calculated using a spatial proximity analysis function in GIS to create a point distance matrix.

Finally, each artifact must be assigned to a (7) *raw material* class. Raw material classifications are user-defined, but for best results, should produce multiple, mutually exclusive groups. The classifications used should depend on the variability of the raw material present at a site. For example, if a site has a variety of raw material types, it may be useful to subdivide by stone type alone (e.g., obsidian, basalt, quartzite, etc.). But, if a site were composed of a single stone type (e.g., chert), subdivisions by stone type alone would not produce discriminating results within the framework of the model. In the latter case, it would be useful to subdivide raw material types into classes that correspond to minimum analytical nodules (MANs). MANs are groupings made by an analyst according to visible similarities in raw material, color, inclusions, texture, grain size, cortex, and fluorescence [22,31–33,35] and can be made *within* a single raw material type. Grouping artifacts into MANs is a common initial step in most refitting projects [9,22,33,35]. (In future versions of the program, separate variables for stone type and MANs will be included, given that many sites contain multiple raw materials, each of which are also marked by considerable nodule variability. In the test case, MAN divisions were sufficient, since a single raw material, quartzite, makes up ~98% of the assemblage.) Regardless of whether raw materials are grouped by stone type or MANs, unique identifiers are assigned to each raw material class and recorded for individual artifacts.

At present, the model is designed to refit debitage, though the model could potentially identify refits to minimally retouched flake tools and utilized flakes. The model was not constructed to identify refits of bifacially retouched tool fragments, such as projectile points. The reason the model is unable to identify refits of this nature is because the variables pertinent to identifying debitage refits are different from those pertinent to identifying projectile point refits. Projectile point refits could be better identified using variables such as point width or thickness, fracture type (e.g., perverse or transverse), flaking style (e.g., parallel oblique or irregular), fragment type (e.g., stem or tip), etc. These are not included in the present version of the model.

In the model, artifacts can refit in two ways. The first is a *surface refit* in which the ventral surface of a flake refits to the dorsal surface from which it was removed. In effect, it involves the refitting of production sequences [16]. The second refit type is the *end-break refit* [16]. In this type of refit, a single broken flake is mended together (e.g., refitting the proximal and distal ends of a single flake). End-breaks could result from breakage during manufacture, intentional snaps, or fracture after deposition. While both the surface refit and end-break refit may result in a successful match, the processes that originally created the artifact pair are different in each case. In an end-break refit, the pair should be composed of two broken flakes. Neither member of the pair should be a complete flake, because in an end-break refit, fragments of a single artifact are united. This is not required of a surface refit, where two separate flakes can be joined regardless of their condition (complete or broken).

As such, separate refitting rules were established for each refit type. These rules were used to create two suitability models: *Model 1* (surface refits) and *Model 2* (end-break refits). The rules that drive the refitting suitability model are outlined in Table 2. For each model, two types of variables are evaluated: those that eliminate unsuitable refits (preliminary screening variables) and those that determine the relative ranking of potentially suitable refits (suitability variables). Raw material serves as a preliminary screening variable in both *Model 1* and *Model 2*, on the assumption that in order for two items to refit, they must have come from the same original raw material group (in this case, the same MAN). An artifact derived

Table 2
Rules used to create ranked list of potential refits

Model 1: surface refit
(A) Rules using preliminary screening variables
(1) Select items in same raw material group
(B) Rules using suitability variables:
(1) Spatially proximal artifacts assigned higher score
(2) Artifacts with same cortex measure assigned higher score
(3) Artifacts with same flake scar count assigned higher score
(4) Artifacts of similar size assigned higher score
Model 2: end-break refit
(A) Rules using preliminary screening variables
(1) Select items in same raw material group
(2) Complete artifacts cannot refit to broken artifacts
(3) Only logical portion matches permitted
(B) Rules using suitability variables
(1) Spatially proximal artifacts assigned higher score
(2) Artifacts with same cortex measure assigned higher score
(3) Artifacts with same flake scar count assigned higher score

from a different raw material group should be classified as an unsuitable refit. Additional preliminary screening variables such as condition and portion apply only to *Model 2*. Broken artifacts cannot refit to complete artifacts in an end-break refit. The portion of the target artifact should determine the other artifacts to which it can refit. For example, proximal flakes should refit to medial, distal, or unspecified flakes, but never to proximal or complete flakes.

Suitability variables include Euclidian distance separating artifacts, percent cortex, flake scar count, and size. Distance can be relevant in determining the probability of a refit because barring post-depositional disturbance, artifacts originating from the same knapping source should be spatially proximal [29,30,42]. The latter three variables are indicative of manufacture stage. It is argued that an artifact has a higher likelihood of refitting to an artifact of a similar manufacture stage. The size variable does not apply to *Model 2* because flake breakage can result in artifacts of any size independent of manufacture stage.

Because each variable is measured in different units of incomparable scales, they must be standardized prior to analysis. Preliminary screening variables are converted to a Boolean rating of 0 or 1 while suitability variables are converted to a common ordinal rating scale ranging from 1 to 9. Ratio variables (e.g., scar count, size, and distance) were recorded as such during analysis, and then converted to ordinal values to produce ratings. The final rating scheme is arbitrary. The schemes used to generate ratings for *Model 1* and *Model 2* are defined in Table 3.

Variable ratings are combined according to the refitting rules in order to obtain a score for each potential refit. The equations used to calculate the final scores are shown in Table 4 where preliminary screening variable ratings are included as multiplicative factors and suitability variable ratings as additive variables. Because each target artifact is defined by a unique combination of attributes, the rating values for input variables vary according to the specific traits of the target artifact. The model dynamically assigns a rating to all artifacts in the refitting population relative to the target artifact. For example, if Target Artifact *X* has two dorsal flake scars, then according to the refitting rules, it should have a greater likelihood of refitting to artifacts with a similar scar count. Therefore, artifacts in the refitting population that have exactly two dorsal flake scars receive a high rating (9), while others receive lower ratings (1, 3 or 6).

Not all suitability variables necessarily play an equal role in the refitting process. Thus, an analyst may feel, given the specific conditions of the site, that distance is a more informative variable than artifact size in finding appropriate refits. Suitability variables can be assigned different weights according to their relative importance. However, determining the relative importance of each variable may not be an easy task, especially as the number of variables increases. Therefore, the model provides a tool to calculate weights using the analytic hierarchy process (AHP) [13,47,48].

AHP is a method to determine the relative priority of multiple criteria where weights are derived through a series of

Table 3
Model 1 and *Model 2* rating schemes

Variable	Value range		Rating
(a) *Model 1* rating scheme			
Raw material		Raw material$_t$ = raw material$_r$	1
		Raw material$_t$ ≠ raw material$_r$	0
Distance[a]		Distance$_{t-r}$ (0–0.5 m)	9
		Distance$_{t-r}$ (0.5–1.0 m)	8
		Distance$_{t-r}$ (1.0–1.5 m)	7
		Distance$_{t-r}$ (1.5–2.0 m)	6
		Distance$_{t-r}$ (2.0–2.5 m)	5
		Distance$_{t-r}$ (2.5–3.0 m)	4
		Distance$_{t-r}$ (3.0–3.5 m)	3
		Distance$_{t-r}$ (3.5–4.0 m)	2
		Distance$_{t-r}$ (4.0–4.5 m)	1
Cortex	0% = 1	Cortex$_t$ = cortex$_r$	9
	1–49% = 2	Cortex$_t$ = cortex$_r$ ± 1	6
	50–99% = 3	Cortex$_t$ = cortex$_r$ ± 2	3
	100% = 4	Cortex$_t$ = cortex$_r$ ± 3	1
Scar count	0 scars = 1	Scar count$_t$ = scar count$_r$	9
	1 scars = 2	Scar count$_t$ = scar count$_r$ ± 1	6
	2 scars = 3	Scar count$_t$ = scar count$_r$ ± 2	3
	3 or more scars = 4	scar count$_t$ = scar count$_r$ ± 3	1
Size	0–19 mm = 1	Size$_t$ = size$_r$	9
	20–39 mm = 2	Size$_t$ = size$_r$ ± 1	5
	40 mm or more = 3	Size$_t$ = size$_r$ ± 2	1
(b) *Model 2* rating scheme			
Raw material	A–S	Raw material$_t$ = raw material$_r$	1
		Raw material$_t$ ≠ raw material$_r$	0
Condition	Complete = 1	Condition$_r$ = 2	1
	Broken = 2	Condition$_r$ = 1	0
Portion		Portion$_t$ or portion$_r$ = CO	0
	Complete = CO	Portion$_t$ = PR, portion$_r$ = PR	0
	Proximal = PR	Portion$_t$ = DS, portion$_r$ = DS	0
	Medial = MS	Portion$_t$ = PR, portion$_r$ = MS, DS, or US	1
	Distal = DS		
	Unspecified = US	Portion$_t$ = DS, portion$_r$ = MS, PR, or US	1
		Portion$_t$ = MS, portion$_r$ = MS, PR, DS or US	1
		Portion$_t$ = US, portion$_r$ = MS, PR, DS, or US	1
Distance[a]		Distance$_{t-r}$ (0–0.5 m)	9
		Distance$_{t-r}$ (0.5–1.0 m)	8
		Distance$_{t-r}$ (1.0–1.5 m)	7
		Distance$_{t-r}$ (1.5–2.0 m)	6
		Distance$_{t-r}$ (2.0–2.5 m)	5
		Distance$_{t-r}$ (2.5–3.0 m)	4
		Distance$_{t-r}$ (3.0–3.5 m)	3
		Distance$_{t-r}$ (3.5–4.0 m)	2
		Distance$_{t-r}$ (4.0–4.5 m)	1
Cortex	0% = 1	Cortex$_t$ = cortex$_r$	9
	1–49% = 2	Cortex$_t$ = cortex$_r$ ± 1	6
	50–99% = 3	Cortex$_t$ = cortex$_r$ ± 2	3
	100% = 4	Cortex$_t$ = cortex$_r$ ± 3	1
Scar count	0 scars = 1	Scar count$_t$ = scar count$_r$	9
	1 scars = 2	Scar count$_t$ = scar count$_r$ ± 1	6
	2 scars = 3	Scar count$_t$ = scar count$_r$ ± 2	3
	3 or more scars = 4	Scar count$_t$ = scar count$_r$ ± 3	1

t = target artifact.
r = artifact in refitting population.
[a] Euclidian distance equation used to find distance $((x_t - x_r)^2 + (y_t - y_r)^2)^{1/2}$.

Table 4
Equations used to calculate suitability score for each artifact in the refitting population

Model 1

$$\text{Score}_r = \text{material}_r[w_1(\text{distance}) + w_2(\text{percent cortex}_r) + w_3(\text{scar count}_r) + w_4(\text{size}_r)]$$

Model 2

$$\text{Score}_r = (\text{material}_r)(\text{condition}_r)(\text{portion}_r)[w_1(\text{distance}) + w_2(\text{percent cortex}_r) + w_3(\text{scar count}_r)]$$

r = artifact in refitting population.
$w_{1\ldots4}$ = assigned weight.

pair-wise comparisons. Rather than having to simultaneously evaluate the importance of multiple variables, with AHP the decision-maker must only judge relative priority between a pair of variables. Relative importance is assigned to pairs according to a standardized system of importance intensity (Table 5). A value of 9 indicates extreme importance of one variable compared to another while a value of 1 indicates equal importance of those variables. The reciprocals of those values represent the corresponding degree of relative unimportance [13,48]. Importance values are input interactively by the analyst into a matrix of pair-wise comparisons within the refitting suitability model. Using an external dynamic link library (EignUtl.dll) through the ''getEigenvectors'' function [37], eigenvectors and eigenvalues can be calculated for the AHP matrix. Standardized weights are obtained for each variable using the eigenvector of the largest eigenvalue [13]. The computer automatically populates the weights for all suitability variables with the derived values.

Once the weights are assigned to each suitability variable and a target artifact is selected, the program calculates potential refits, and generates a ranked list of potential refits in table form. This list contains the artifact identifier and its corresponding score calculated from the equation. The score for each artifact estimates its likelihood of refitting to the target artifact, where a higher score indicates a greater likelihood. Scores range from 9 (most likely to refit) to 1 (least likely to refit), while a score of 0 indicates an impossible refit. For each target artifact, the number of potential refits varies and does so according to the number of artifacts eliminated during preliminary screening. The list of ranked potential refits can then be used as a guide to help find actual refits, where the highest ranked artifact is the first artifact one should attempt to refit to the target artifact, and so on iteratively.

Table 5
Importance intensity values for AHP pair-wise comparison matrix

Value	Definition
1	Equal importance
3	Moderate importance
5	Strong importance
7	Very strong or demonstrated importance
9	Extremely important
2, 4, 6, 8	For compromise between above values
Reciprocals of above values	If variable *i* has one of the above non-zero numbers assigned to it when compared to activity *j*, then *j* has the reciprocal value when compared with *i*
Rationals	Ratios arriving from the scale
1.1–1.9	For tied variables

Table adapted from Charnpratheep *et al.* [13].

3.2. Testing the refitting suitability model

In order to test the accuracy of this refitting suitability model, an archaeological control sample of artifacts with previously identified refits were input into the program. The goal was to determine whether the model accurately replicated known refits. As a secondary goal, the test procedure aimed to identify possible weaknesses of the model and thus provide useful information for improving the technique in the future.

All refit pairs ($n = 89$) came from 5GN149, a high elevation surface lithic scatter in Gunnison County, Colorado, USA. The artifacts consist primarily of debitage, and extend over an area of $\sim$3500 m^2. Several high-density hotspots have been identified. Research has focused on Cluster 1 (Fig. 1), a well-defined artifact scatter covering roughly 12 m^2. Cluster 1 contains approximately 1900 artifacts, all of which were individually piece-plotted. Preliminary analyses suggest that the chipping cluster may represent a relatively discrete flint-knapping episode. As such, 5GN149 was viewed as an ideal candidate for refitting. In order to establish controls for testing the refitting suitability model, refits were initially identified through the process of traditional refitting (Fig. 1).

While most refitting analyses do not use spatial proximity as a variable (i.e., the distance between two artifacts does not affect whether or not the analyst attempts to refit them), the refitting suitability model factors distance when identifying and ranking suitable refits. Percussion flaking experiments suggest that artifacts within lithic scatters are not randomly distributed, but instead their dispersal is patterned [29,30,42]. Moreover, artifacts from the same flint-knapping episode disperse from their percussion source an average of less than 1 m, and routinely not more than about 5 m. Therefore, it can be expected that artifacts originating from the same knapping episode should be found closer to one another than artifacts originating from a different knapping episode. In most cases, an artifact should have a higher probability of refitting to an artifact in close proximity than to an artifact separated by a considerable distance. Knapping locales that have not been significantly disturbed after deposition should demonstrate a similar spatial relationship. Given its well-defined boundary, Cluster 1 does not appear to have undergone substantial post-depositional transformation, and therefore spatial proximity is included as a variable in this test of the refitting suitability model. In those cases where the analyst wishes to test for contemporaneity among spatially disparate clusters, the spatial variable will need to be formulated differently.

For a given refit pair from the 5GN149 assemblage, each artifact was input separately into the model as the target artifact. Each target artifact was then compared against only the artifacts from Cluster 1 that fell into its same distinct raw material class (MAN). Artifacts that could not be confidently grouped by MAN were excluded. The total control sample comprised 1370 artifacts and served as the pool of potential refits. The list of potential refits generated for one artifact in

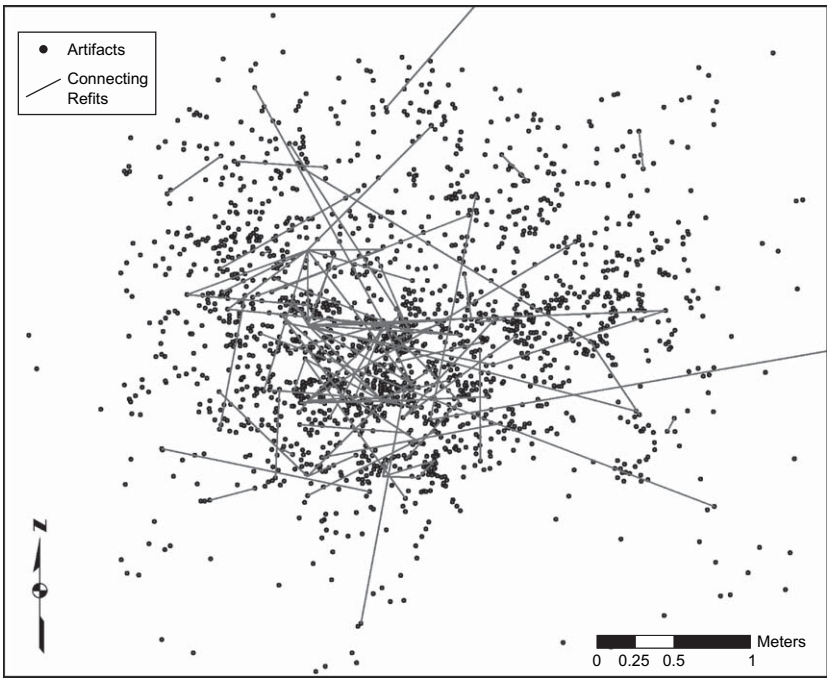

Fig. 1. Plan map of Cluster 1 from 5GN149 in Gunnison County, Colorado, USA. Cluster 1 contains 1900 pieces of stone chipping debris, all of which were piece-plotted using an EDM Total Station. Lines connect refit artifacts. The refit line that extends off the northern edge of the map is 3.4 m long; the line off the eastern edge is 14.0 m long.

a pair is not necessarily complimentary to the list generated for its mate. This is so because each artifact in a refit pair is distinguished by a particular combination of characteristics, and the list of potential refits is customized to that specific artifact. Therefore, unless two artifacts within a refit pair have identical characteristics (including spatial position), their list of potential refits will differ.

The control sample of target artifacts includes 89 refit pairs. However, each artifact within a pair is treated as an independent case, resulting in 178 cases by which the model could be tested. *Model 1* (surface refits) and *Model 2* (end-break refits) were tested separately in order to assess the reliability of each model. In the control sample, 72 artifacts belonged to surface refits and 106 artifacts belonged to end-break refits.

Two test trials were conducted (Table 6). In Trial 1, all variables received equal weight (*Model 1*: all weights = 0.25; *Model 2*: all weights = 0.333). In Trial 2, AHP was used to derive the weights (Table 7). In the paired comparisons within the AHP matrix, the distance criterion was assigned dominant priority (Table 6). If the results of Trial 2 prove more accurate than Trial 1, then it can be argued that proximity is an important factor in determining refits at 5GN149. If the results of Trial 2 prove less accurate than Trial 1, then distance is no more influential than other variables in determining refits. These results can then help us expand our understanding of site formation at 5GN149.

4. Results

In order to gauge the success of the refitting suitability model, the model had to be evaluated relative to the alternative approach, traditional refitting without benefit of suitability assessment. As mentioned earlier, the results of any traditional refitting project are largely contingent on the skill of the refitting analyst, making it difficult to model the success rate of a traditional refitting project, which would vary depending on the skill of the refitting analyst. Therefore, we attempted to approximate the refitting success that might be expected from an inexperienced analyst, who, in order to be systematic and thorough, would identify refits through a series of pairwise comparisons. In this approach, each artifact has an equal opportunity of being examined by the analyst. This is, in effect, the minimal success rate one might expect from the traditional approach.

The complete results of Trial 1 are summarized in Fig. 2. Percentages are used to describe the rank of known refits, because for each target artifact the model reduces through preliminary screening the total artifact pool ($n = 1370$) to a smaller, but variable, sample of potential refits. Because each target artifact has a unique combination of traits, the subsequent sample of potential refits varies in constituency and

Table 6
Weight assignments for 5GN149 test trials

	Model 1: surface refit	Model 2: end-break refit
Trial 1: without AHP	All variables receive equal weight (0.25)	All variables receive equal weight (0.333)
Trial 2: with AHP	Distance: 0.5158 Cortex: 0.1894 Flake scars: 0.1894 Length: 0.1054	Distance: 0.6 Cortex: 0.2 Flake scars: 0.2

Table 7
Value matrix used to calculate weights for Trial 2 using AHP

	Distance	Cortex	Flake scars	Length
Model 1				
Distance	1	3	3	4
Cortex	0.33	1	1	2
Flake scars	0.33	1	1	2
Length	0.25	0.5	0.5	1

	Distance	Cortex	Flake scars
Model 2			
Distance	1	3	3
Cortex	0.33	1	1
Flake scars	0.33	1	1

size. The remaining artifacts are then ranked according to their suitability of refitting to the target artifact. Therefore, in order to compare model success between all target artifacts used in this pilot study, the rank of known refits is referred to as a standardized percentage rather than numerical value.

Based on the combined results of *Model 1* and *Model 2*, the refitting suitability model placed the known refit in the top ranking group (1–10) approximately 32 of the time (Table 8). Known refits were identified in the lower and lowest ranking groups (i.e., 40–50,…,80–90) significantly less often. By correctly placing the known refit in the top group of potential refits approximately 32% of the time, the refitting suitability model yielded a higher success rate than would be expected from simple pair-wise comparisons, where a refit has only a 10% chance of being identified within the first 10 refit attempts, assuming the likelihood of finding a successful refit is random. Therefore, given the results above, the refitting suitability model identified 22 out of 100 more refits in the top ranking group (1–10) than would be expected through random

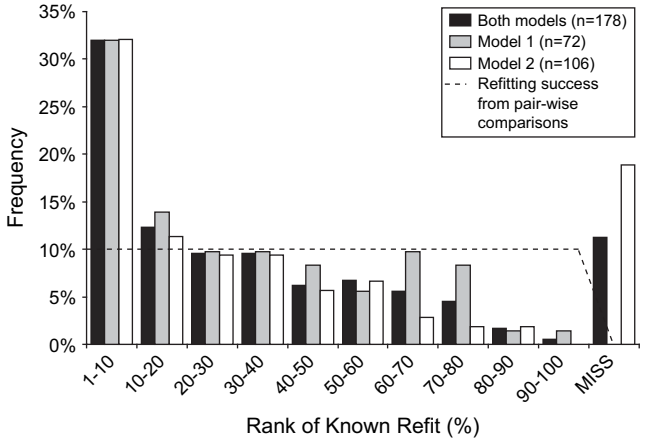

Fig. 2. Results of Trial 1. Figure shows combined results of *Models 1* and *2*, as well as individual results for each model. Solid bars indicate the frequency at which the refitting suitability model identified the known refit within each ranking group. The higher the frequency of known refits in the leftmost categories, the better, because it indicates that the refitting suitability model correctly identified the known refit at the top of the list of potential refits. The dashed line indicates the expected frequency distribution if the process of identifying refits was performed through pair-wise comparisons.

Table 8
The frequency at which known refits were identified in the top ranking group (0–10) by the GIS refitting model

	Model 1: surface refit (%)	Model 2: end-break refit (%)	Models 1 and 2 combined (%)
Trial 1: without AHP	31.94	32.08	32.02
Trial 2: with AHP	20.83	33.96	28.65

pair-wise comparisons. Overall, the observed results using the refitting suitability model differs significantly from the expected results of pair-wise comparisons ($G = 69.097$, $p < 0.001$, where G measures the goodness of fit between an observed and expected frequency distribution; based on data in Fig. 2).

However, the refitting suitability model also failed to include the known refit in the list of ranked refits for 20 of the 178 target artifacts. A missed refit occurred when the model eliminated the known refit during the preliminary screening process, and assigned it a score of 0 and therefore excluded it from the pool of potential refits. All of the missed refits occurred when the *Model 2* equation was applied. None of the missed refits occurred with the *Model 1* equation. This suggested that the *condition* and *portion* variables, both used in the *Model 2* but not *Model 1* equation, were likely responsible for eliminating the actual refits.

In order to identify more specifically the circumstances by which known refits were eliminated, each missed refit was individually inspected. There were two circumstances under which missed refits occurred: first, when either the target artifact or the known refit was mistakenly coded as a complete artifact; and second, when the relationship between the target artifact portion and the known refit portion was illogical (e.g., proximal to proximal), and thus considered by the program to be impossible to refit. Thus, in both instances, missed refits were the result of misidentification of the artifact during attribute analysis, rather than a structural flaw of the suitability model.

Fig. 3 compares the accuracy of Trial 1 to Trial 2. In Trial 1 all variables received equal weight. In Trial 2, AHP was used to determine weights, with greater relative importance assigned to the distance variable. When the results from *Model 1* and *Model 2* are grouped, Trial 2 produced less accurate results than Trial 1 overall. The refitting program ranked known refits in the top ranking group (1–10) less often than in Trial 1, implying that distance may be no more important than other variables in determining refits. However, when the results of *Model 1* and *Model 2* are viewed separately, a slightly different pattern is revealed. In Trial 2, *Model 1* performed noticeably worse. This suggests that for surface refits, distance plays a less important role in determining whether two artifacts refit in the 5GN149 assemblage. On the other hand, *Model 2* performed slightly better in Trial 2, though the distributions in both trials are similar. That the *Model 2* results improved when distance was assigned greater weight relative to other variables may suggest that end-break refit pairs are more likely to involve spatially proximal artifacts. This is not surprising, given that end-break refits join segments which once comprised

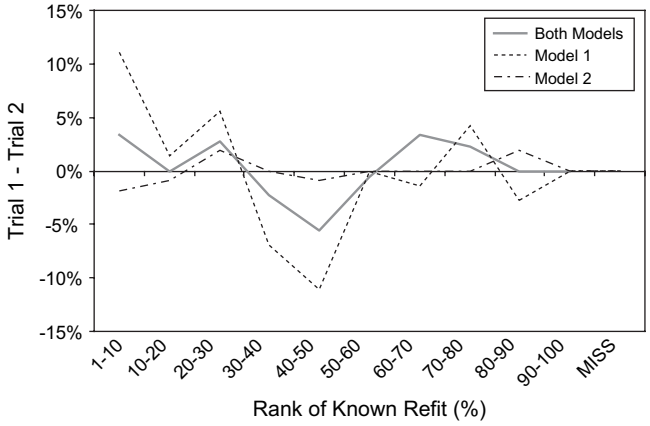

Fig. 3. The compared success rates of Trial 1 (all variables receive equal weight) and Trial 2 (weights assigned through AHP, where distance receives higher priority). When the line falls above the *x*-axis for a given ranking group, Trial 1 results placed more known refits in that group. When the line falls below the *x*-axis for a given ranking group, Trial 2 results placed more known refits in that group.

a single artifact. If the artifact broke during manufacture or post-deposition, it might still be located close to its refit.

In sum, the individual and combined results for *Model 1* and *Model 2* for both Trial 1 and Trial 2 of this pilot study demonstrate the refitting suitability model's ability to assign high rank to known refits (Table 8). Again, the model is able to identify known refits in the top ranking group (1−10) more frequently than expected through pair-wise comparisons alone, where only 10% of known refits would be identified in the top ranking group.

5. Discussion

The goals of the refitting suitability model are twofold. The first goal is to reduce the time and labor required by refitting in order to make the method easier for archaeologists; the second is to standardize and systematize the refitting process so as to increase the comparability of analysis results. If these goals are met in conjunction, then the refitting process might be used more frequently in archaeological analysis.

Whether or not the refitting suitability model reduces the time and labor of refitting is complicated to assess. The model has been shown to successfully identify known refits significantly better than would be expected by random pair-wise matching alone and is therefore an efficient alternative to that specific approach. For a novice refitting analyst, the model offers a helpful tool to guide the refitting process. Yet, for an experienced refitting analyst, who would not refit using pair-wise comparisons, the refitting suitability model may only present a minimal advantage. With the model, skilled and unskilled analysts alike should be able to achieve the same results and in an equal amount of time.

In fact, for either the experienced or inexperienced refitting analyst, the refitting suitability model offers a convenient starting point for a refitting project. Moreover, since refitting is a cumulative process where the discovery of a single refit often leads to the discovery of others [9], to have a starting sample of completed refit pairs might be valuable to the analyst.

A clear strength of the refitting suitability model is its potential to standardize refitting so that results of different refitting projects can be compared. By using the model, refitting analysts evaluate a series of common and clearly defined variables. The process of assessing potential refits is not conducted idiosyncratically, but instead is performed systematically by the model. An analyst can record and report the variables and weights that he/she used for a specific assemblage in order to ensure the comparability of different assemblages. In sum, it attempts to remove some of the subjectivity of the process.

Under certain conditions, the refitting suitability model may not be the appropriate choice. One potential obstacle of the model is that it requires initial acquisition of the requisite attribute data. If a researcher intends to collect these data for other reasons, using the refitting suitability model adds no analytical cost. But, with mass-analysis procedures becoming more common in lithic studies, some researchers may no longer record detailed attribute data for individual artifacts within assemblages [1,23]. In the latter case, refitting without the suitability model may be cost-effective, because coding attributes for each individual artifact would require additional time. Also, at present, the refitting suitability model requires that each artifact have a tight spatial provenience. Piece-plotted data are best suited for the refitting suitability model, though these data can be very time-consuming to acquire. Investigators may have a difficult time justifying this level of precision at all sites [43,57]. Of course, other versions of this model can be developed which do not require piece-plotted data.

Program results are shaped entirely by the attributes assigned to an artifact during analysis. If an item is misclassified (e.g., a proximal flake is coded as a complete flake), which is an ever-present possibility given the inaccuracies and inconsistencies in lithic analysis [5,18], then a refit might be missed, as occurred in the 5GN149 test case. Therefore, until our identification skills are further honed, the refitting suitability model has the potential to produce erroneous results, since the model itself cannot correct for identification error. However, even though the refitting suitability model can result in missed refits, traditional refitting can as well. While the traditional approach does not eliminate artifacts in the process of preliminary screening, as does the refitting suitability model, in traditional refitting, missed refits can result from failure of the analyst to (1) attempt a refit between two artifacts or (2) recognize a refit upon attempt. Only under specific (and unrealistic) conditions, including unlimited analysis time and flawless recognition skills, would an analyst find all possible refits in an assemblage. Therefore, even though the refitting suitability model missed a small percentage of the known refits, a certain degree of error would be expected with the traditional refitting method as well. Further, because the model has the potential to expedite refitting, it may lead to a more sophisticated understanding of tool manufacture, thereby improving lithic analysis and reducing attribute recognition error. This feedback process could in turn result in improved performance of the model.

The effectiveness of the refitting suitability model may vary in accordance with site context because the model in its current form emphasizes the spatial proximity of refittable artifacts. At present, the model is particularly appropriate for use on isolated knapping areas. At sites where artifact distributions have been significantly altered by anthropogenic or post-depositional processes, refittable artifacts may no longer follow a predictable spatial patterning. For example, if in the past a person removed an artifact from a lithic scatter and used it as a tool elsewhere in the site, the model might not assign the artifact a high ranking. Also, the refitting suitability model might not account for phenomena like artifact "toss zones," indicated in the ethnographic record [8]. But, these are some of the processes archaeologists seek to detect, because they speak to past human behavior. Although this particular test and iteration of the model is biased against that particular portion of the behavior archaeologists hope to identify, further research will explore tests of the model in which spatial data are deemphasized or disregarded in order to permit recognition of long-distance refits.

The results of any refitting program must be taken cautiously. Even if two items refit, contemporaneity is not necessarily established [34]. Repeated use of sites and recycling of raw materials can produce misleading patterns, linking two unrelated behaviors [24,27,34,41]. Also, surface sites that have undergone considerable post-depositional transformation may produce spurious refitting results if non-cultural factors contributed to artifact patterning [7,40]. Refitting also differentially emphasizes the importance of expedient tool use because those tools tend to be used at their manufacture location [34]. Curated tools, by definition, are removed from their original manufacture context [34] and would not be found amongst refits.

The proposed refitting suitability model alone cannot resolve the problems associated with refitting studies. But, it offers a way to make the process easier for archaeologists. It also standardizes and systematizes the method, and thus increases the comparability of separate refitting projects. As such, the model helps to improve the analytical benefits of refitting and reduces the time and labor costs of the method.

Future enhancements to the model could result in a higher success rate for the model and thus further reduce the cost–benefit ratio. These might include the removal or addition of variables in order to improve refitting efficiency and accuracy. The relative importance of a variable can be assessed by systematically adjusting its weight in test trials. Variables that consistently reduce refitting success can then be eliminated from the model. Also, the rating system might be altered so that variables are subdivided into a greater number of categories in order to produce more discriminating results. Multiple permutations of the model should be tested in order to achieve maximal success.

The refitting suitability model should also be further tested using known refits from archaeological contexts other than 5GN149. Future trials will help to identify the conditions under which the model performs optimally. These data will enable an analyst to make an informed decision on whether or not the refitting suitability model is an appropriate tool for his/her own analysis.

At present, the refitting suitability model has not been used to identify novel refits—that is, refits that have not already been recognized through traditional refitting. This will be the next step in testing the model.

6. Concluding remarks

Preliminary results suggest that the refitting suitability model offers an effective alternative approach to refitting. In this pilot study, approximately 32% of the time the known refit was assigned high rank (1–10) and therefore placed at the top of the list of potential refits. These results are significantly better than would be expected by iterative random matching alone.

The refitting suitability model can potentially make refitting easier for archaeologists. As such, archaeologists can apply refitting more often and to a wider range of problems. The under-investigated lithic scatter can thus receive greater attention. This may result in an improved understanding of prehistoric stone tool technology and of spatial patterns in the archaeological record. The model also offers a more systematic and standardized approach to refitting. All variables and weights used during refitting are explicitly defined, offering better control over the process. In addition, the model will produce the same results regardless of the skill of the refitting analyst. Therefore, if the variables and weights are maintained between analyses, the results of different refitting projects can be compared and be used to make meaningful interpretations about a site or lithic assemblage. The differences between the results of two refitting projects can be attributed to differences in past behaviors or site taphonomy, rather than variation in the skill of the refitting analyst or the time devoted to the process.

Hofman [25] predicted that "[i]n time, refitting studies will probably become standard practice in analyses of chipped-stone assemblages given certain contextual settings and research questions." Nearly 25 years later, refitting studies are far from standard, especially in North American archaeology. Given the demonstrable analytical value of refitting, this is unfortunate. Indeed, with lithic scatters comprising the majority of the archaeological record, we should hope to see refitting as one of the more common tools used by archaeologists. That refitting is so infrequently used suggests we may be overlooking a significant amount of information in the archaeological record, reducing the analytical potential of archaeological sites. The refitting suitability model proposed here can help expedite and standardize the process, thereby making refitting a more applicable and powerful analytical tool in addressing archaeological problems.

Acknowledgements

This project was made possible by the financial support of the QUEST Archaeological Research Program at Southern Methodist University. We would like to acknowledge the National Park Service for granting access to 5GN149 under

permit to D.J. Meltzer and B.N. Andrews (NPS Antiquities Permit 03CURE1), with thanks to F. Frost and A. Anderson for their assistance. We are grateful for all of the help received throughout the entirety of this project, especially from field and lab crew-members B. Andrews, R. Byerly, H. Cooper, S. Cooper, C. Dukeman, A. Graves, J. Hargrave, T. Jennings, D. Meltzer, B. Mueller, J. Roberson, J. Taylor-Montoya, and C. Wolff. We are appreciative of the ideas contributed by T. Rick and G. Sampson. We are also grateful to R. Briggs and B. Berry for their helpful comments on an earlier version of this paper. We would like to especially thank D. Meltzer for his guidance, suggestions, and encouragement along the way. We would also like to acknowledge M. Aldenderfer, the discussant for the SAA session in which an earlier version of this paper was presented. In addition, we are grateful to L. Jennings for her assistance on statistical problems and J. Huang for his programming contributions. Finally, we thank M.L. Larson and an anonymous reviewer for their helpful comments and suggestions on an earlier draft of this paper.

References

[1] S.A. Ahler, Mass analysis of flaking debris: studying the forest rather than the trees, in: D.O. Henry, G.H. Odell (Eds.), Alternative Approaches to Lithic Analysis (1989), pp. 85–118 Archaeological Papers of the American Anthropological Association No. 1, Washington, DC.

[2] S.A. Ahler, Use-phase classification and manufacturing technology in Plains Village arrowpoints, in: J.L. Hofman, J.G. Enloe (Eds.), Piecing Together the Past: Applications of Refitting Studies in Archaeology, BAR International Series 578, Oxford, 1992, pp. 36–62.

[3] W. Andrefsky Jr., Lithics: Macroscopic Approaches to Analysis, Cambridge University Press, Cambridge, 1998.

[4] R.N.E. Barton, C.A. Bergman, Hunters at Hengistbury: some evidence from experimental archaeology, World Archaeology 14 (1982) 237–348.

[5] C. Beck, G.T. Jones, Bias and archaeological classification, American Antiquity 54 (1989) 244–262.

[6] C.A. Bergman, Effective research and analysis at upland sites Byways to the Past IV Conference Proceedings (2002). <http://www.pennbyways.net/Byways2002cd/Proceedings/Session%20II/Bergman%20with%20links.pdf> (accessed July 2005).

[7] C.A. Bergman, J.F. Doershuk, How the data come together: refitting in lithic analysis, Journal of Middle Atlantic Archaeology 8 (1992) 139–160.

[8] L.R. Binford, In Pursuit of the Past: Decoding the Archaeological Record, Thames and Hudson, New York, 1982.

[9] P. Bleed, Refitting as aggregate analysis, in: C.T. Hall, M.L. Larson (Eds.), Aggregate Analysis in Chipped Stone, University of Utah Press, Salt Lake City, 2004, pp. 184–198.

[10] C.A. Bollong, Analysis of site stratigraphy and formation processes using patterns of pottery sherd dispersion, Journal of Field Archaeology 21 (1994) 15–28.

[11] D. Cahen, Refitting stone artifacts: why bother? in: G. De G. Sieveking, M.H. Newcomer (Eds.), The Human Uses of Flint and Chert Cambridge University Press, Cambridge, 1987, pp. 1–9.

[12] D. Cahen, L.H. Keeley, F.L. Van Noten, Stone tools, toolkits, and human behavior in prehistory, Current Anthropology 20 (1979) 661–683.

[13] K. Charnpratheep, Q. Zhou, B. Garner, Preliminary landfill site screening using fuzzy geographical information systems, Waste Management and Research 15 (1997) 197–215.

[14] A.E. Close, Reconstructing movement in prehistory, Journal of Archaeological Method and Theory 7 (2000) 49–77.

[15] S.N. Collcutt, N.R.E. Barton, C.A. Bergman, Refitting in context: a taphonomic case study from a Late Upper Palaeolithic site in sands on Hengistbury Head, Dorset, Great Britain, in: E. Cziesla, S. Eickhoff, N. Arts, D. Winter (Eds.), The Big Puzzle: International Symposium on Refitting Stone Artefacts, Studies in Modern Archaeology, vol. 1, Holos, Bonn, 1990, pp. 219–235.

[16] E. Cziesla, On refitting of stone artifacts, in: E. Cziesla, S. Eickhoff, N. Arts, D. Winter (Eds.), The Big Puzzle: International Symposium on Refitting Stone Artefacts, Studies in Modern Archaeology, vol. 1, Holos, Bonn, 1990, pp. 9–44.

[17] E. Cziesla, S. Eickhoff, N. Arts, D. Winter (Eds.), The Big Puzzle: International Symposium on Refitting Stone Artefacts, Studies in Modern Archaeology, vol. 1, Holos, Bonn, 1990.

[18] C.D. Dukeman, An analysis of analysts: a critical assessment of the replicability of debitage attributes, unpublished MA thesis, Department of Anthropology, University of Wyoming, Laramie, 2002.

[19] S. Eickhoff, A spatial analysis of refitted flint artefacts from the Magdalenian site of Gönnersdorf, Western Germany, in: E. Cziesla, S. Eickhoff, N. Arts, D. Winter (Eds.), The Big Puzzle: International Symposium on Refitting Stone Artefacts, Studies in Modern Archaeology, vol. 1, Holos, Bonn, 1990, pp. 307–330.

[20] H. Floss, T. Terberger, The Magdalenian of Andernach: analysis of camp structures by refitting stone artifacts, in: E. Cziesla, S. Eickhoff, N. Arts, D. Winter (Eds.), The Big Puzzle: International Symposium on Refitting Stone Artefacts, Studies in Modern Archaeology, vol. 1, Holos, Bonn, 1990, pp. 339–362.

[21] J.I. Gross. Some patterns of chipping activity of the Thunderbird Clovis component, in: W.M. Gardner (Ed.), The Flint Run Paleoindian Complex: A Preliminary Report 1971–1973 Seasons, Occasional Paper Number 1, Department of Anthropology, Catholic University of America, Washington, DC, 1974, pp. 100–106.

[22] C.T. Hall, Evaluating prehistoric hunter–gatherer mobility, land use, and technological organization strategies using minimum analytical nodule analysis, in: C.T. Hall, M.L. Larson (Eds.), Aggregate Analysis in Chipped Stone, University of Utah Press, Salt Lake City, 2004, pp. 139–155.

[23] C.T. Hall, M.L. Larson (Eds.), Aggregate Analysis in Chipped Stone, University of Utah Press, Salt Lake City, 2004.

[24] B. Hayden, Comments on Cahen et al. Current Anthropology 20 (1979) 673–674.

[25] J.L. Hofman, The refitting of chipped-stone artifacts as an analytical and interpretive tool, Current Anthropology 22 (1981) 691–693.

[26] J.L. Hofman, Defining buried occupation surfaces in terrace sediments, in: J.L. Hofman, J.G. Enloe (Eds.), Piecing Together the Past: Applications of Refitting Studies in Archaeology, BAR International Series 578, Oxford, 1992, pp. 128–150.

[27] J.L. Hofman, Putting the pieces together: an introduction to refitting, in: J.L. Hofman, J.G. Enloe (Eds.), Piecing Together the Past: Applications of Refitting Studies in Archaeology, BAR International Series 578, Oxford, 1992, pp. 1–20.

[28] M.A. Jodry, Fitting together Folsom: refitted lithics and site formation processes at Stewart's Cattle Guard site, in: J.L. Hofman, J.G. Enloe (Eds.), Piecing Together the Past: Applications of Refitting Studies in Archaeology, BAR International Series 578, Oxford, 1992, pp. 179–209.

[29] K.L. Kvamme, Investigating chipping debris scatters: GIS as an analytical engine, in: D.G. Maschner (Ed.), New Methods, Old Problems: Geographic Information in Modern Archaeological Research, Occasional Paper, No. 23, Center for Archaeological Investigations, Southern Illinois University, Carbondale, 1996, pp. 38–71.

[30] K.L. Kvamme, Patterns and models of debitage dispersal in percussion flaking, Lithic Technology 22 (1997) 122–138.

[31] M.L. Larson, Towards a holistic analysis of chipped stone assemblages, in: P.J. Carr (Ed.), The Organization of North American Prehistoric Chipped Stone Tool Technologies. International Monographs in Prehistory, Archaeological Series 7, Ann Arbor, 1994, pp. 57–69.

[32] M.L. Larson, Chipped stone aggregate analysis in archaeology, in: C.T. Hall, M.L. Larson (Eds.), Aggregate Analysis in Chipped Stone, University of Utah Press, Salt Lake City, 2004, pp. 3–17.

[33] M.L. Larson, J.B. Finley, Seeing the trees but missing the forest: production sequences and multiple linear regression, in: C.T. Hall, M.L. Larson

(Eds.), Aggregate Analysis in Chipped Stone, University of Utah Press, Salt Lake City, 2004, pp. 95–111.
[34] M.L. Larson, E.E. Ingbar, Perspectives on refitting: critique and a complementary approach, in: J.L. Hofman, J.G. Enloe (Eds.), Piecing Together the Past: Applications of Refitting Studies in Archaeology, BAR International Series 578, Oxford, 1992, pp. 151–162.
[35] M.L. Larson, M. Kornfeld, Chipped stone nodules: theory, method, and examples, Lithic Technology 22 (1997) 4–18.
[36] J. Lindenbeck, Sitefit-3D-refitting with the PC: a preview, in: E. Cziesla, S. Eickhoff, N. Arts, D. Winter (Eds.), The Big Puzzle: International Symposium on Refitting Stone Artefacts, Studies in Modern Archaeology, vol. 1, Holos, Bonn, 1990, pp. 73–80.
[37] O. Marinoni, Implementation of the analytical hierarchy process with VBA in ArcGIS, Computers & Geosciences 30 (2004) 637–646.
[38] A.E. Marks, P. Volkman, Changing core reduction strategies: a technological shift from the Middle to Upper Paleolithic transition in the Levant, in: E. Trinkaus (Ed.), The Mousterian Legacy: Human Biocultural Change in the Upper Pleistocene, BAR International Series 164, Oxford, 1983, pp. 13–33.
[39] A.E. Marks, P. Volkman, Technological variability and change seen through core reconstruction, in: G. de G. Sieveking, M.H. Newcomer (Eds.), The Human Uses of Flint and Chert, Cambridge University Press, Cambridge, 1987, pp. 11–20.
[40] G.C. Mohapatra, Comments on Cahen et al. Current Anthropology 20 (1979) 674.
[41] T.M. Morrow, Refitting and archaeological site formation processes: a case study from the Twin Ditch Site, Greene County, Illinois, in: G.H. Odell (Ed.), Stone Tools: Theoretical Insights into Human History, Plenum Press, New York, 1996, pp. 345–373.
[42] M.H. Newcomer, G. de G. Sieveking, Experimental flake scatter-patterns: a new interpretative technique, Journal of Field Archaeology 7 (1980) 345–352.
[43] J.F. O'Connell, Ethnoarchaeology needs a general theory of behaviour, Journal of Archaeological Research 3 (1995) 205–255.
[44] J. Riel-Salvatore, M. Bae, G.A. Clark, J.M. Lindly, P. McCartney, A. Razdan, The past meets the future: 3D modeling technology and lithic analysis at Wadi al-Hasa locality 623X, Journal of Human Evolution 42 (2004) A29.
[45] J. Riel-Salvatore, M. Bae, P. McCartney, A. Razdan, Paleolithic archaeology and 3D visualization technology: recent developments, Antiquity 76 (2002) 929–930.
[46] D.C. Roper, Lateral displacement of artifacts due to plowing, American Antiquity 41 (1976) 372–375.
[47] T.L. Saaty, The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation, McGraw-Hill, New York, 1980.
[48] T.L. Saaty, Highlights and critical points in the theory and application of the analytic hierarchy process, European Journal of Operational Research 74 (1994) 426–447.
[49] J. Schäfer, Conjoining artefacts and consideration of raw-material: their application at the Middle Paleolithic site of the Schweinskopf-Karmelenberg, in: E. Cziesla, S. Eickhoff, N. Arts, D. Winter (Eds.), The Big Puzzle: International Symposium on Refitting Stone Artefacts, Studies in Modern Archaeology, vol. 1, Holos, Bonn, 1990, pp. 83–100.
[50] E. Schaller-Åhrberg, Refitting as a method to separate mixed sites: a test with unexpected results, in: E. Cziesla, S. Eickhoff, N. Arts, D. Winter (Eds.), The Big Puzzle: International Symposium on Refitting Stone Artefacts, Studies in Modern Archaeology, vol. 1, Holos, Bonn, 1990, pp. 611–622.
[51] U. Schurmans, A. Razdan, A. Simon, P. McCartney, M. Marzke, D. van Alfen, G. Jones, J. Rowe, G. Farin, D. Collins, M. Zhu, D. Liu, M. Bae. Advances in geometric modeling and feature extraction on pots, rocks and bones for representation and query via the Internet, in: In Archaeological Informatics: Pushing The Envelope CAA2001. Computer Applications and Quantitative Methods in Archaeology, BAR International Series 1016 (2002), Archaeopress, Oxford.
[52] M.F. Seeman, Intercluster lithic patterning at Nobles Pond: a case for "disembedded" procurement among early Paleoindian societies, American Antiquity 59 (1994) 273–288.
[53] J.F. Simek, The organization of lithic technology and evolution: notes from the continent, in: P.J. Carr (Ed.), The Organization of North American Prehistoric Chipped Stone Tool Technologies, International Monographs in Prehistory, Archaeological Series 7, Ann Arbor, 1994, pp. 118–122.
[54] C.A. Singer, The 63-kilometer refit, in: J.E. Ericson, B.A. Purdy (Eds.), Prehistoric Quarries and Lithic Production, Cambridge University Press, Cambridge, 1984, pp. 35–48.
[55] W.G. Smith, Man the Primeval Savage, Edward Stanford, London, 1894.
[56] P. Spikins, Virtual landscapes: GIS and lithic scatters, in: J.A. Schofield (Ed.), Lithics in Context. Lithic Studies Society Occasional Papers No. 5, Lithic Studies Society, 1995, pp. 95–104.
[57] P. Spikins, H. Ayestaran, C. Conneller, Meaningful patterns of mere pictures: the influence of collection strategy on the interpretation of lithic scatters, in: J.A. Schofield (Ed.), Lithics in Context. Lithic Studies Society Occasional Papers No. 5, Lithic Studies Society, 1995, pp. 125–137.
[58] F.C.J. Spurrell, On the discovery of the place where Paleolithic implements were made at Crayford, Quarterly Journal of the Geological Society 36 (1880) 544–549.
[59] S. Veil, A dynamic model of a Magdalenian settlement by spatial analysis of refitted artifacts, in: E. Cziesla, S. Eickhoff, N. Arts, D. Winter (Eds.), The Big Puzzle: International Symposium on Refitting Stone Artefacts, Studies in Modern Archaeology, vol. 1, Holos, Bonn, 1990, pp. 45–60.
[60] P. Villa, Conjoinable pieces and site formation processes, American Antiquity 47 (1982) 276–290.
[61] M. Wansleeben, Fitting refitting into a computer, in: E. Cziesla, S. Eickhoff, N. Arts, D. Winter (Eds.), The Big Puzzle: International Symposium on Refitting Stone Artefacts, Studies in Modern Archaeology, vol. 1, Holos, Bonn, 1990, pp. 63–72.
[62] D.G. Wyckoff, Refitting and protohistoric knapping behavior: the Lowrance example, in: J.L. Hofman, J.G. Enloe (Eds.), Piecing Together the Past: Applications of Refitting Studies in Archaeology, BAR International Series 578, Oxford, 1992, pp. 83–127.