

Conceptual premises in experimental design and their bearing on the use of analogy: an example from experiments on cut marks

Manuel Domínguez-Rodrigo

Abstract

Experimental archaeology embodies a large array of conceptual approaches. In the present work, it is argued that only those methodological approaches maximizing comparability between experiments and case-specific archaeological problems are heuristically scientific. This reduces the range of analogies that can be applied to the past. The adequateness of analogies depends on how the conceptual premises of experiments are designed. A practical example of this is provided through the comparison of referential frameworks created to understand the utility of cut marks to reconstruct butchering behaviors.

Keywords

Analogy; assumptions; premises; hypothesis testing; cut marks.

Introduction

Experimental archaeology and actualistic research are integral parts of ‘middle-range’ theory and thus of modern scientific archaeology, which is based on the testing of alternative hypotheses (Binford 1981; Gifford 1981). Hypotheses are framed within specific referential analogs created by careful observation. These referential frameworks are elaborated by controlled documentation of processes in which behaviors of independent agents are understood within their specific contexts and the resulting actions of these agents are diagnosed (Binford 1981; Gifford 1981; Gifford-González 1991; Gould 1965, 1979; Wylie 1982, 1988, 1989).

If there is a hierarchy of principles that can be applied to the components of actualistic research, it can be argued that the most important one is the adequate use of premises (see Wylie 1988) in the elaboration of referential frameworks. Researchers create these analogs primarily to understand behaviors represented in and responsible for the archaeological record. The significance of analogy as a non-objective entity was initially stressed by Richter (1928). It entails a series of assumptions, some of them selected by the researcher, in a dialectic dynamic between the ideas that researchers try to test and the way the testing is eventually carried out.

A systemic evolutionary taphonomic approach (innovated by Fernández-López 2006), considering taphonomic entities as endowed with properties subjected to change according to their structure, behavior and environment, also shows that the selection of criteria to be replicated in experiments is ultimately dependent on what has been called taphonomic redundancy.¹

A widely accepted articulation of theoretical principles guiding actualism was outlined by Gifford-Gonzalez (1991). She differentiated between 'formal' analogy and 'relational' analogy. The former is obtained through observation and the latter through inference. She conceived of a continuum from one form of analogy to the other within a hierarchical conception of taphonomic processes defined by six nested analytical categories (trace, causal agent, effector, actor, behavioral context and ecological context). Formal analogies can be used in the first four categories, since actors can be observed, provided equifinality can be overcome (Lyman 2004). In contrast, the behaviors and the ecological factors that determine them can never be directly reconstructed from the analysis of bones and have to be indirectly inferred. In this case, relational analogy applies. Gifford-Gonzalez (1991) argued that the six analytical categories were interdependent. Starting from the broadest categories, Gifford-Gonzalez argued that every single taphonomic process is understood primarily in specific ecological contexts. If ecology conditions behavior, then actors should react in a predictable way according to those conditions; in turn, any such actions should be reflected in the traces imprinted on bones. Any experiment that obviates the relationship of these nested categories would be conceptually flawed.

Every analog is in essence incomplete, since it reproduces only a selected and limited set of variables, and can control for only a determined number of these. Similar processes in the past where other non-experimentally considered variables may have intervened make the application of analogs systematically imperfect. Given that researchers must be aware of the imperfect nature of analogy, the relevance of the correct use of premises and assumptions in experiment design cannot be overemphasized.

Some analogies in taphonomy can be defined as substantive since they reproduce general processes that are not subject to a significant degree of variability. For instance, the patterns of bone breakage (notches, planes) resulting from experimenting with physical processes such as dynamic (hammerstone) or static (carnivore dentition) loading are more generally applicable as analogs than other processes subject to greater contextual variability. In many studies involving controlled experimentation of physical processes limited to the actor-trace sequence, analogies can justifiably be used within generalized referential frameworks. In contrast, and more specifically in archaeology, analogies depending on ecological-behavioral factors are subject to a higher degree of variability and can be confidently used as referential frameworks only of determined taphonomic

problems; they are case-specific, and could be labeled methodological analogies. In this type of analogy, given the large array of variables at play, researchers have to be aware of the list of assumptions they are making, of how these assumptions translate into hypothesis premises (Wylie 1988) and, eventually, how these premises and the hypotheses that contain them are subjected to testing. Failure to do so will produce false equifinality scenarios and ambiguity in interpretation and eventually fuel post-processual criticism of the subjective nature of the scientific method.

The present work will use experimental studies of cut marks as an example of the variability of criteria used by researchers when conducting experiments and designing referential frameworks. It will be argued that this variability is not always scientifically acceptable, either because some approaches to experiment design are conceptually flawed (incoherent use of assumptions and premises of what is supposed to be replicated) or, in other cases, because comparisons across experimental datasets cannot be sustained when the premises of the hypotheses tested by different researchers are not the same.

Analogy, uniformitarianism and the concept of regularity

The only way to ‘reconstruct’ the past is to assume that there are certain regularities in the way the world works which are not subject to time, and are therefore observable in the present. Thus, these regularities can also be inferred for the past. The assumptions of uniform rates and the implication of slow and gradual change in ‘substantive uniformitarianism’, using Gould’s (1965) term, have been proved incorrect in many cases. The modern conception of uniformitarianism does not assume the constant rate of change and acknowledges that the agents of change cannot be verified empirically. However, the laws that govern these agents remain permanent. This new uniformitarianism is methodological and vital to scientific procedure. Spatial and temporal invariability in the laws that control processes is absolutely critical if any general conclusion about the past is to be made from observations in the present (Gould 1965). Methodological uniformitarianism does not directly inform on nature, but provides an approach with which to understand it (Shea 1982). This approach, in assuming that natural laws are invariant in time and space, does not invoke unknown hypothetical processes if the observed results can be explained through modern processes (Gould 1965).²

Simpson (1970) further elaborated on this concept by separating those aspects of the world that remain unmodified in time and space (immanent properties) and those that are contingent on particular interactions in each moment and place (configurational properties). Immanent properties allow historical processes (or parts of them) to be interpreted precisely because they are not subject to variation in space and time, i.e. they are universal. Configurational properties must be approached in a different way. Unlike universal immanent properties, configurational properties are based on *regularities* in the variables that regulate them. It is important to note, however, that there are two types of configurational processes: (1) those that are highly variable and therefore difficult to predict; and (2) those that are highly regular and therefore predictable. Obviously, only the latter can be reliably applied to past dynamics.

I agree with Gould (1980) in that only those processes whose properties and range of variation can be measured should be used in scientific archaeology. Some researchers believe that only geological/physical processes can be understood from such an approach (e.g. Nairn 1965). However, Simpson (1970) stresses that for any process to be understood it needs only to be uniform; that is, it needs to exhibit regular properties. This is possible whether reconstructing the mechanic aspects of the world or biotic behaviors. Much misunderstanding emanates from the misconception that only universal, or immanent, laws are applicable to the past. However, because even universal laws are never absolute (Popper 1956, 1972), we are left with heuristic explanations (Lakatos 1978) that are grounded in the predictability of their regularities.

Regularity is derived from probability. However, Simpson (1970) notes two difficulties with inferring historical processes: (1) multiple processes may have similar results (that is, equifinality); and (2) configuration makes processes unpredictable. Simpson (1970) himself suggests a solution for the latter. Scientific prediction depends on what is periodical and repetitive. Although historical events are unique and therefore unpredictable on at least some level, there are different degrees of distinctiveness, and historical events can be considered predictable to the extent to which we understand their causes and the regularity of their behavior. This is how Simpson (1970) defines historical laws: based on probabilities and with similar heuristic power (*sensu* Lakatos 1978) as natural laws. An historical event is determined by the immanent characteristics of the universe that act upon it, but in a configurational way.

This leads us to the use of analogy. To understand past configurational processes, the context and variables that generate modern processes must also be understood. This understanding differentiates descriptive from formal analogies and from relational (dialectical) analogies (Gifford-Gonzalez 1991). Analogies play an important role because they can discern and document variability in observed regularities. Most importantly, analogies can be observed and replicated. To use an analogy properly as a referential framework (*sensu* Binford 1981) for interpreting taphonomic processes in the past, researchers must clearly be able to: (1) differentiate whether the analogy is case-specific or general; (2) in the former case, produce a list of assumptions using collected data from the assemblage where hypothesis testing will take place; (3) justify that the experimental premises match the set of assumptions made for the formulation of a hypothesis. This can be better explained with an example.

A practical example documenting conceptual variability in hypothesis testing: experimental replication and interpretation of cut marks

The use of replication in experimental archaeology during the 1980s enabled a certain optimism that cut marks could be scientifically used to infer human butchery behaviors (Binford 1978, 1981; Bunn 1981, 1982; Bunn and Kroll 1986; Gifford 1977; Lyman 1987; Gifford-González 1989). Nowhere has this been better illustrated than in the hunting-scavenging debate concerning Plio-Pleistocene sites in East Africa (for a review of this debate, see Domínguez-Rodrigo 2002). The possibility that cut marks could be linked equally with hunting and with scavenging behaviors prompted the development of new

experimental protocols of opposite-hypothesis testing to distinguish both behaviors (Domínguez-Rodrigo 1997a, 1997b). Nilssen (2000) also contributed with new experimental protocols to differentiate diverse butchery behaviors. However, in the past ten years, the diverse array of experimentation on carcass butchery has yielded a varied interpretative repertoire comprising the following claims (Table 1):

1. Cut marks are of limited value to interpret butchery behaviors and the order of access to carcasses by hominids because they are subject to equifinality, given that they could be the result of the removal of the scraps of flesh surviving carnivore consumption of their prey (Capaldo 1995, 1997, 1998).
2. Cut marks, when applied to early Plio-Pleistocene sites, support the hypothesis that they were the result of hominids butchering carnivore kills (Selvaggio 1994).
3. Cut-mark patterns found in Plio-Pleistocene sites, when compared to those documented in modern foragers (Hadza), support a mixed strategy of early, intermediate and late access to variously fleshed carcasses (Lupo and O'Connell 2002).
4. Actualistic referential frameworks are useful to interpret cut marks as resulting from primary access to fleshed carcasses by humans *versus* de-fleshed carcasses abandoned by carnivores, and support the hypothesis of primary access to fleshed

Table 1 Interpretation of the methodological utility of cut marks and of the data on these bone surface modifications from east African Plio-Pleistocene sites by the main researchers whose research is discussed in the present work

<i>Researchers</i>	<i>Methodological interpretation of cut marks</i>	<i>Archaeological interpretation of Plio-Pleistocene hominid carcass acquisition and butchery</i>
Selvaggio (1994)	Subject to equifinality.	Hominids were passive scavengers from carnivore kills.
Capaldo (1995, 1997, 1998)	Subject to equifinality.	Hominids were passive scavengers from felid kills and mass drownings.
Domínguez-Rodrigo (1997a, 1997b)	Cut-mark frequencies and anatomical distribution (per element and bone section) can be used to differentiate between butchery of fleshed carcasses (primary access) and removal of scraps of flesh from some carnivore kills (secondary access).	Hominids had primary access (probably including hunting) as the dominant strategy of carcass acquisition.
Lupo and O'Connell (2002)	Not valid to discriminate primary versus secondary access to carcasses since they are statistically indistinguishable in both experimental scenarios.	Hominids used mixed strategies of hunting and confrontational scavenging (in various stages of carcass completeness).
Pobiner (2007)	Subject to equifinality.	Hominids were passive scavengers from felid kills.

- carcasses by hominids (Domínguez-Rodrigo 1997a, 1997b, 2002; Domínguez-Rodrigo and Pickering 2003).
5. Comparisons of different experimental sets of cut marks, which test primary or secondary access to carcasses, show that they do not provide resolution, since they are statistically indistinguishable (Lupo and O'Connell 2002; see critique in Domínguez-Rodrigo 2003).
 6. Recent experiments have widened the degree of variability of flesh abandoned in carnivore kills, which suggests that cut-mark patterns previously derived from experiments that recorded flesh distribution in a more restrictive sample of kills are no longer valid, providing more evidence of the behavioral ambiguity of cut marks (Pobiner 2007).
 7. The tremendous range of variation in frequencies and anatomical distributions of cut marks across multiple assemblages prompts skepticism that the behavioral meaning of cut marks could be effectively inferred from prehistoric assemblages (Lyman 2005).

The obvious message is this: the ambiguity of the cut mark data limits their power to explain butchery behaviors and, therefore, the order of access to carcasses by hominids and carnivores. Most of the experiments, and interpretation of cut marks, in the points listed above (six out of seven) have been carried out and applied to a restricted number of Plio-Pleistocene sites in East Africa in order to understand the butchery behavior that these sites have preserved for our interpretation of a crucial stage of human evolution. More specifically, most have been applied to one site: FLK Zinj. This clearly shows that, in principle, the focus of these experiments was case-specific. However, in blatant contradiction to this, most statements on the meaning of cut marks (this author's included) were thought to be of universal applicability. This is wrong: cut-mark frequency and anatomical distribution are the result of processes that belong to the 'ecological' and 'behavioral' spheres of Gifford-Gonzalez's (1991) nested set of inferences and are, therefore, subjected to variability. This prevents any experiment carried out to test the meaning of cut marks in the kind of 'inferred' ecological and behavioral contexts to be applied anywhere else where both variables may have been different. As Lyman recently admitted, 'well-founded interpretations of frequencies of cut-marked remains may require unique kinds of contextual data' (2005: 1722). FLK Zinj was formed in an alluvial 'near-lake' habitat within an ecosystem where felids and hyenas seem to have been fairly abundant (Domínguez-Rodrigo et al. 2007). Given that resource availability for scavenging hominids is ecologically dependent (Blumenschine 1986; Tappen 1992), when modeling opposing hypotheses of access to carcasses, researchers have to elaborate their experimental premises to try to ensure maximum consistency between experiments and the context inferred.

I will use the example of the behavioral meaning of cut marks at FLK Zinj to illustrate how a specific set of assumptions, premises and hypotheses can be designed and successfully tested. I will also use a comparison with currently available experimentation to explain why some researchers may be closer than others to accurately testing the hypotheses of primary or secondary access to carcasses by hominids and, indirectly, to providing high-resolution (rather than ambiguous) referential frameworks. This

comparative exercise can be graphically followed in Figure 1. The null hypothesis is that cut marks lack sufficient resolution from which to infer primary or secondary access to carcasses. A subsequent null hypothesis is that hominids were scavengers (secondary access hypothesis). Proving that both versions of the null hypothesis are wrong would imply that early hominids had primary access to carcasses and that this can be inferred by specific placement and frequencies of cut marks.

Assumption 1 The essence of any experimental study is control. The only way to link actor-effector-causal agent-trace effectively is by having as much control as possible of the complete experimental/observational process. In the case of the hypotheses under testing, one factor in which control is key is resource availability from carnivore kills as potential scavengeable resources for hominids. This is especially relevant in the case of flesh scraps. The assumption is that no data derived from uncontrolled experiments should be heuristically used in this regard, because we may be inferring the wrong actor, producing an equivocal diagnosis. The resulting premise is that the experiment has to be carried out with as much control as possible or otherwise discarded.

For the secondary access hypothesis, the experiments that are inadequate according to this premise are those made by Pobiner (2007) in the wild, who *never* witnessed a complete

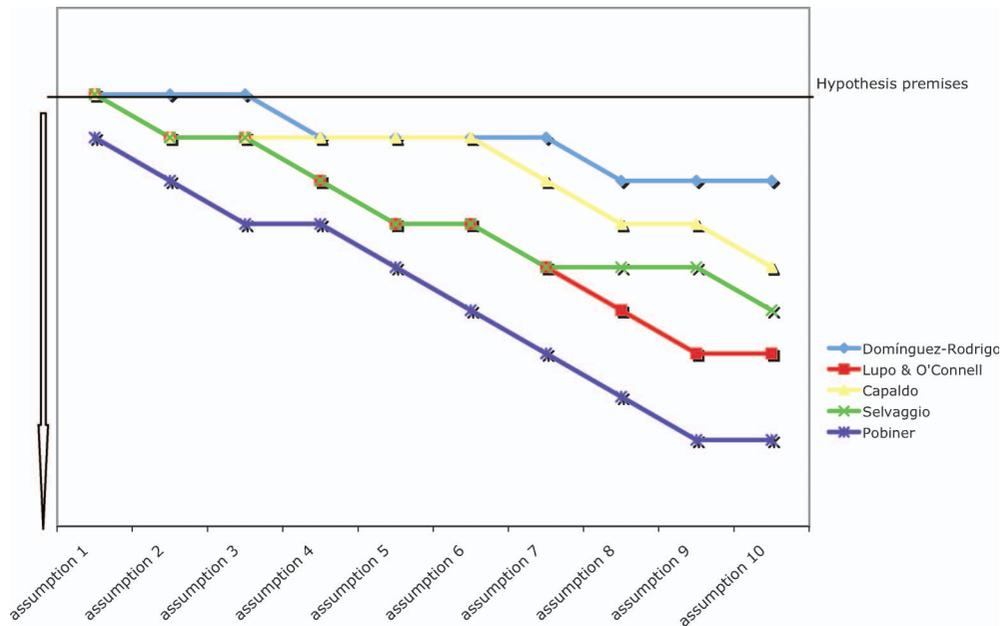


Figure 1 Experimental matrix showing the conceptual assumptions and premises of hypothesis testing as described in the text (top horizontal line) and the deviations from these by various researchers. Each experimental premise for each researcher that does not take any given outlined assumption into consideration is reflected in a step down from the horizontal line of the matrix. The lower the experimental research appears (direction of arrow) compared to the top horizontal line, the more conceptually distant the experiment is from the hypothesis premises, and the more inappropriate it is for comparison with the specific case of the behavioral meaning of cut marks at FLK Zinj.

process of carcass consumption by her lion sample, since she documented hunts in the late evening-early night and evaluated carcass modification and resource availability the next morning. Her study lacks control and is based on inferences that cannot be empirically supported. The reported tooth-mark damage from the carcasses that she collected in the wild may also be the result of other carnivores having access to carcass remains during the night. This could explain why the only controlled sample that she collected in captivity shows a more intense consumption of flesh than that reported in her wild 'lion' sample and, in apparent contradiction, almost half the tooth-mark frequency.

Assumption 2 Following a basic Popperian principle, hypotheses can be tested only when confronted with their opposite. Our whole understanding of the use of statistics in science is based on this principle: the null hypothesis. Experiments used to test a hypothesis have to be able to test the opposite and reject it. Inferences drawn from unilateral testing are not scientifically reliable. The assumption made is that an opposite-testing hypothesis is well founded only when the same set of assumptions, premises and analytical variables has been used. This happens most frequently when it is the same researcher who carries out the testing of both hypotheses. The scientific premise is that only in equally comparable analytical sets can opposite hypotheses be tested and compared.

In the comparative set of experiments, most researchers have unilaterally tested a hypothesis, relying on the results obtained by a different researcher for the opposite hypothesis. However, since the set of variables used by every researcher is unique (see description in Domínguez-Rodrigo 2003), the comparisons are not necessarily valid.

Assumption 3 The constraints of elementary taphonomic alteration (as defined by Fernández-López 2006) are determined primarily by the ecological context where it takes place. FLK Zinj was formed in a near-lacustrine habitat where a large array of carnivores was present. Actualistic studies have shown that the interplay between felids and hyenids is the most determinant for understanding modern bone modification and deposition and resource availability for scavengers in modern African savannas (Blumenschine 1986; Domínguez-Rodrigo 1996; Tappen 1992). Competition conditions the way each carnivore consumes prey remains. It also conditions the way that terrestrial felids feed. When under pressure from either hyenas or human, felids tend to consume their carcasses hastily, leaving more scraps of flesh (Domínguez-Rodrigo 1999). At Olduvai, during both Bed I and Bed II times, hyenas seem to have been using the alluvial habitats with even higher frequency than in similar modern settings (Domínguez-Rodrigo et al. 2007; Monahan 1996). This can be inferred from the intensity of hyenid-modified assemblages in these settings with no modern equivalence. This has specific relevance regarding the amount of scavengeable resources by hominids. The assumption and subsequent premise made from this inference is that experimental replication of cut marks has to be carried out (especially those on carcasses obtained from carnivore kills) in similarly competitive settings to guarantee comparability.

When applied to the comparative experimental set (Fig. 1), all researchers but one comply with this premise. Capaldo (1995, 1998) and Selvaggio (1994) made their experiments in the Serengeti. Domínguez-Rodrigo (1997a) carried out his experiments in Maasai Mara, Tsavo, Galana and Kulalu. Lupo and O'Connell (2002) made their observations in Eyasi – with a much lower presence of carnivores than the national parks where the previous authors carried out their research, but similarly diverse in carnivore

taxa. In contrast, Pobiner (2007) conducted her research on a Kenyan private ranch, where some carnivores were systematically hunted. Most of the hyenas were either poisoned or shot at and, during the years that she used the ranch for her research, most of the predators were lions; she documented the presence of only two to five leopards and hardly any cheetah. Given their abundance, lions were also shot sometimes, prompting them to be mostly nocturnal (L. Frank pers. comm.). In this human-altered ecosystem, lion behavior was conditioned by two variables: marginal inter-specific competition due to the removal of hyenas and the human impact on the demographics of lions. As a result, the amount and anatomical distribution of flesh that Pobiner documented in carcasses abandoned by lions differ (in some cases drastically) from the more consistent descriptions reported by Selvaggio (1994) and Blumenschine (1986) for the Serengeti and Ngorongoro ecosystems and Domínguez-Rodrigo (1997a, 1997b, 1997c, 1999) for the Maasai Mara, Tsavo, Galana and Kulalu ecosystems, which were more similar among themselves. Domínguez-Rodrigo et al. (2007) report that, in the Galana and Kulalu ecosystems, lions showed a double pattern of carcass consumption: one affecting wild game, in which they consumed carcasses similarly to those documented in national parks, abandoning only very marginal scraps of flesh, and one affecting domestic animals (cattle), which they hunted during the night, transporting them away from human camps and abandoning them before dawn after incomplete consumption. Thus, it can be seen that diverse environments in different ecosystems in national parks with minimal anthropic impact yield very different results from those documented in a highly altered environment, such as the one used by Pobiner (2007), or in specific situations of human-lion interactions as documented by Domínguez-Rodrigo. This supports the claim that only experiments conducted in environments unmodified by humans reliably document variability in carnivore behavior. Indeed, while Tsavo, Galana and Kulalu offered different ecological conditions from Maasai Mara, Serengeti and Ngorongoro, the documented way of lion consumption of wild game was very similar in resulting flesh availability.

Pobiner's (2007) study obviates the ecological impact of the altered environment where she conducted her study and she claims that the results obtained are heuristically useful for discriminating the real utility of cut marks inferred from the amount and variation in the anatomical distribution of flesh abandoned by felids. From the experimental frameworks currently available for understanding flesh abandoned by felids in the Zinj environment, Pobiner's is the least appropriate given the drastic ecological differences documented between the two types of environment.³

Assumption 4 Adequateness of the sample. Experimental samples should replicate (as much as possible) the archaeological samples in terms of the range of animal size and the range of body parts represented. This can be further defined by two independent analytical variables described by Domínguez-Rodrigo (2003): animal size used in butchery experiments (small versus large) and experiment type (using complete carcasses, all limbs or only a few limb bones). Carcasses accumulated at Zinj comprise a large number of individuals, documented (despite the abundance of limb bones) by all skeletal elements from small and large animals. The assumption is that an experiment replicating complete carcass consumption of small and large individuals would more accurately reflect what happened at Zinj than experiments based on a few bones from a single carcass size and from a single individual. Domínguez-Rodrigo and Barba (2005) and Pobiner and Braun

(2005) showed that cut-mark patterns could be distinct in different carcass sizes. The premise is that, to maintain the appropriateness of the comparability of experiments, cut-mark patterns obtained from specific carcass sizes should not be applied to interpret cut marks in different carcass sizes. Also, in the Zinj case, the use of complete carcasses for experimentation may be more adequate than partial carcasses. Only Capaldo (1995) uses this premise correctly. The other researchers either use one variable alone correctly, or both variables are inadequate. While Pobiner (2007) used complete carcasses, she never conducted a thorough study of cut marks using these variables and involving defleshing and demarrowing of bones as documented archaeologically.

Assumption 5 Sample size and composition (see discussion in Domínguez-Rodrigo 2003) are crucial for correct inference. Sample sizes in all the compared sets of experiments are highly variable, from large samples, like those obtained by Capaldo (1995, 1998) or Domínguez-Rodrigo (1997a, 1997b) for butchery of fleshed carcasses, to samples composed of multiple experiments on single elements or a pair of bones per carcass (Selvaggio 1994), which do not reproduce the assumption that carcasses were accumulated at Zinj in a more complete state (whether hunted or scavenged). The interpretative model developed by Pobiner (2007) for flesh availability on large carcasses at lion kills is derived from a total of nine carcasses in contrast with Domínguez-Rodrigo's (1997c, 1999) sample of twenty-nine individuals where flesh distribution was documented and almost twenty carcasses from lion kills where secondary access was experimentally modeled. If an arbitrary threshold of a minimum of ten carcasses⁴ per hypothesis tested (comprising at least complete limbs in each experiment) is used as a premise, some researchers' samples would be left out (Fig. 1).

Assumption 6 In order to interpret the validity of cut marks for inferring different butchery behaviors and primary or secondary access to carcasses by hominids, the observation of the anatomical distribution of flesh (whether bulk or scraps) in carnivore kills is not enough; experimental butchery is also necessary. Within Gifford-Gonzalez's (1991) conceptual scheme of hierarchical order of inference, the documentation of flesh at carnivore kills would be situated in the ecological sphere. From there to the final obtainment of cut-mark patterns (traces), one should be able to document how actors, with the aid of effectors, produce specific traces. Pobiner (2007) did not conduct any of these experiments and her interpretation of the use of cut marks rests on the assumption that one can skip the experimental process linking ecology and traces by indirect assumption. This is conceptually flawed.

Assumption 7 Butchery at Zinj was carried out with stone tools. The primary access hypothesis would assume that, if hominids were hunters, they had regular access to carcasses and they may have been efficient and knowledgeable butchers. The alternative hypothesis, that hominids were scavengers that had only sporadic access to carcass remains, would imply that they may not have been expert butchers. The premise in the former hypothesis is that experimental butchery must be carried out by expert butchers, since butchery implies a learning process that is reflected in the decreasing number of cut marks imparted to the bones according to experience; novice butchers leave more cut marks on bones than expert butchers (Domínguez-Rodrigo 1997c). The premise in the latter hypothesis is that stone tools would be used to remove every single scrap of flesh and would not be focused on flesh bulk removal alone, which is most habitual in common

butchery practices. For both hypotheses, a second premise is that the use of stone tools, preferably of the same raw material type as is archaeologically documented, is an experimental requirement.

Most researchers use these variables (tool type and butcher type) differently. Capaldo's and Lupo and O'Connell's butchery samples were made with metal knives, whereas Selvaggio's and Domínguez-Rodrigo's implied the use of stone tools of the same kind as found at Zinj. However, Selvaggio did not consider the experience of the butcher an important factor, and probably obtained higher frequencies (especially in certain bones) of cut marks than if using an expert butcher.

Assumption 8 The assumption of the type of carcass processing carried out at Zinj depends on the hypothesis to be tested. In the primary access hypothesis, processing assumes three butchery behaviors: skinning, disarticulation and de-fleshing. From an optimal foraging point of view, the secondary access hypothesis assumes that the most efficient behavior is the removal of the flesh scraps at the kill. Disarticulation is the most time-costly butchery activity and also produces the highest degree of tool wear. Disarticulation of fairly de-fleshed carcasses, like those that one would obtain at carnivore kills, is unnecessary. In both cases, it is assumed that de-marrowing followed. The premise is that in each of these hypotheses no other type of processing activity should be experimentally reproduced.

Of the comparative sample of experiments, Domínguez-Rodrigo did not reproduce skinning and disarticulation. Capaldo's introduced an activity (periosteum removal) which is unnecessary for the butchery of most elements. This apparently irrelevant activity can actually produce a high frequency of tool marks on bones, biasing the assumed butchery behavior described above. Lupo and O'Connell's butchered carcass sample obtained from Hadza also included another processing activity not assumed for the Zinj hominids: grease extraction by bone boiling. Whereas this may not directly affect the resulting cut-mark frequencies, it affects them indirectly by limiting the type of bone fragmentation introduced by post-ravaging hyenas, thus, conditioning the resulting frequencies. Hyenas have been suggested to play a secondary role in bone modification and fragmentation at FLK Zinj (Bunn and Kroll 1986; Domínguez-Rodrigo et al. 2007). Bone fragmentation ultimately determines bone surface modification frequencies. The hyena was one of the agents, other than hominids, which probably played a role in bone breakage at Zinj. Experiments suggest that hyenas are interested in modifying bones in human-accumulated bone assemblages when grease is available and preferably while it is fresh (Marean et al. 1992, 2004; Capaldo 1995; Pickering et al. 2003). By deterring hyenas from early access to bones or by removing grease from bones during boiling, hyena post-ravaging is modified and, therefore, the degree of bone fragmentation is also modified, affecting the resulting bone surface modification frequencies.

Assumption 9 The species of the animals used for butchery experiments may also determine both the amount of flesh available for secondary access and the resulting cut-mark pattern from their processing. For instance, equids have stronger muscular attachments to bones (as reflected in the stronger muscular/ligament insertions on certain bones (e.g. caudal tibia and femur) than bovids and the fact that consumption of their bones by carnivores tends to leave more flesh upon abandonment (personal observation).⁵ Likewise, human bulk de-fleshing of equids also produces more scraps of flesh than it does

in bovids. As a result, equids tend to appear more highly cut-marked than bovid remains. An example is provided by Lupo and O'Connell (2002) with various assemblages created by Hadza. In these bone sets, zebras tend to appear cut-marked at rates more than one-third higher than bovids. Most of the processed animals at FLK Zinj were bovids. Therefore, the assumption is that experimental butchery should preferably be carried out on the same kind of carcasses for the sake of comparability. The premise is that experiments should be made by using bovids to test both hypotheses, and that experiments based on either butchery of equids or observation of flesh distribution in carnivore kills composed only of equids are not heuristically valid.

All but one of the experimental samples used for this comparative approach accepted the premise. Pobiner's (2007) sample of lion kills in the wild is primarily composed of zebras (eight zebras and one eland). The eland appears more de-fleshed than several zebras and mid-shafts from upper limb bones appear virtually de-fleshed (see Fig. 2.10 in Pobiner 2007: 56). The resulting pattern documented in the equid sample is not adequate to infer flesh availability in bovids scavenged from lion kills.

Assumption 10 This is one of the most important assumptions. Flesh is differentially distributed across the anatomy of an animal. The assumption is that a methodological approach which does not consider the type of element and the actual location of cut marks on these elements would not accurately reflect human butchery behaviors and the dynamics of these *vis-à-vis* the available scraps of flesh from carnivore kills, which are differentially distributed across the anatomy of carcasses. See an extensive methodological critique in Domínguez-Rodrigo (2002). The premise is that, to reflect butchery behaviors accurately, cut marks have to be tallied according to element and bone section as described in Domínguez-Rodrigo (1997a).

In the experimental comparative sample, Capaldo and Selvaggio are the only ones not to consider this assumption. They use a general method of tallying marks according to bone portion, irrespective of element type and the actual location of marks. They are also the only ones who have experimentally advocated equifinality in the use of cut marks as a result of their method, which lacked resolution in differentiating butchery behaviors.

Discussion

The range of interpretations about the use of cut marks to infer human butchery behaviors, derived from the experimental sets compared in the present work, is not a reflection of the variability of these behaviors and their ecology but is a methodological artifact of the diversity of assumptions made in experimental design and their corresponding experimental premises. Researchers have reacted differently to what is supposed to be tested and the way testing was conducted. Some researchers claim ambiguity in the use of cut marks not because they can prove it, but because of their methods of documenting cut marks (assumption 10) or because they disregarded the determinant interrelated inferential categories of ecology-behavior and skipped the hierarchy of inferential categories. Others failed to document the utility of cut marks to reconstruct butchering behaviors because their selection of premises to articulate their hypotheses and the corresponding variables used during experimentation were different

from those that should have been inferred and used to interpret the targeted fossil assemblage. Figure 1 shows how far each researcher is from the experimental matrix created by the articulation of assumptions and their corresponding premises. The more distant the experimental sets are from the matrix, the less heuristically appropriate they are to interpret cut marks from the fossil assemblage. Some of the studies casting doubt on the utility of cut marks (e.g. Pobiner 2007) mistake the degree of comparability of the datasets used, disregarding ecology, behavioral variability and confiding in untested assumptions to the point of not even replicating butchery when testing the secondary access hypothesis.⁶

Only one of the researchers whose work has been compared claims that cut marks can successfully be used to differentiate between primary access and secondary access to fleshed carcasses (Domínguez-Rodrigo 1997a). Challengers to this claim could support their position either by proving that the set of assumption-premises used by that researcher is equivocal or by using the same experimental premises to document a greater variety of results than that reported in the referential framework provided by his research. Instead of that, by selecting a different experimental path they set themselves up to test-prove something similar but essentially different.

This brings us to reconsider the use of analogy. Given the variability of criteria when designing experiments and when using analogies, an outline containing the set of inferences and premises guiding hypothesis modeling seems necessary. This will help researchers understand when new results from experiments provide compelling new evidence challenging established ideas, or when they simply represent the testing of a completely different set of premises and assumptions, even if the hypotheses appear to be the same.

Acknowledgements

I am most grateful to A. K. Outram for his timely edition of the special volume on experimental archaeology and allowing me to be part of it. I am also grateful to M. Prendergast for her help and support. I appreciate the valuable comments from S. Fernández-López and three anonymous reviewers.

Department of Prehistory, Complutense University

Notes

- 1 Taphonomic redundancy has been defined as the capacity of taphonomic elements to repeat the same message. Taphonomic redundancy, as well as replication, allows the estimation of the 'repeatedness' of taphonomic groups under particular environmental conditions, on the basis of their actual properties (Fernández-López 2006).
- 2 Gould was rewriting Occam's Razor: one should not increase, beyond what is necessary, the number of entities required to explain something. Even fourteenth-century scholars can remind experimental archaeologists of the need to keep it simple.

- 3 This refers to human-modified ecosystems (like that used by Pobiner for her experiments) and savannas not impacted on by humans (like those used by the other researchers referenced in the previous paragraph).
- 4 Experiments with a smaller number of carcasses usually yield large variation ranges, which make hypothesis testing more difficult.
- 5 Several zebras were used by Domínguez-Rodrigo (1997a, 1997b, 1997c, 1999) in his observation of availability of flesh in lion kills and in his experimental replication of the scavenging hypothesis.
- 6 Pobiner (2007) assumes that the distribution of flesh in lion kills that she observed is enough to discredit the utility of cut marks when applied to fossil assemblages, without proceeding to remove flesh through butchery and compare it to a null hypothesis experimental scenario.

References

- Binford, L. R. 1978. *Nunamiut Ethnoarchaeology*. New York: Academic Press.
- Binford, L. R. 1981. *Bones: Ancient Men, Modern Myths*. New York: Academic Press.
- Blumenshine, R. J. 1986. *Early Hominid Scavenging Opportunities: Implications of Carcass Availability in the Serengeti and Ngorongoro Ecosystems*. Oxford: BAR International Series, 283.
- Bunn, H. T. 1981. Archaeological evidence for meat-eating by Plio-Pleistocene hominids from Koobi Fora, Kenya. *Nature*, 291: 574–77.
- Bunn, H. T. 1982. Meat-eating and human evolution: studies on the diet and subsistence patterns of Plio-Pleistocene hominids in East Africa. Doctoral dissertation, University of California, Berkeley.
- Bunn, H. T. and Kroll, E. M. 1986. Systematic butchery by Plio-Pleistocene hominids at Olduvai Gorge, Tanzania. *Current Anthropology*, 27: 431–52.
- Capaldo, S. D. 1995. Inferring hominid and carnivore behavior from dual-patterned archaeological assemblages. Doctoral thesis, Rutgers University, New Brunswick.
- Capaldo, S. D. 1997. Experimental determinations of carcass processing by Plio-Pleistocene hominids and carnivores at FLK 22 (Zinjanthropus), Olduvai Gorge, Tanzania. *Journal of Human Evolution*, 33: 555–97.
- Capaldo, S. D. 1998. Simulating the formation of dual-patterned archaeofaunal assemblages with experimental control simples. *Journal of Archaeological Science*, 35: 311–30.
- Domínguez-Rodrigo, M. 1996. A landscape study of bone conservation in the Galana and Kulalu (Kenya) ecosystem. *Origini*, 20: 17–38.
- Domínguez-Rodrigo, M. 1997a. Meat-eating by early hominids at the FLK 22 Zinjanthropus site, Olduvai Gorge, Tanzania: an experimental approach using cut mark data. *Journal of Human Evolution*, 33: 669–90.
- Domínguez-Rodrigo, M. 1997b. A reassessment of the study of cut mark patterns to infer hominid manipulation of fleshed carcasses at the FLK Zinj 22 site, Olduvai Gorge, Tanzania. *Trabajos de Prehistoria*, 54: 29–42.
- Domínguez-Rodrigo, M. 1997c. Testing meat-eating in early hominids: analysis of cut-marking processes on defleshed carcasses. *Human Evolution*, 12: 169–82.
- Domínguez-Rodrigo, M. 1999. Flesh availability and bone modification in carcasses consumed by lions. *Palaeogeography, Palaeoclimatology and Palaeoecology*, 149: 373–88.

- Domínguez-Rodrigo, M. 2002. Hunting and scavenging by early humans: the state of the debate. *Journal of World Prehistory*, 16: 1–54.
- Domínguez-Rodrigo, M. 2003. On cut marks and statistical inferences: methodological comments on Lupo & O'Connell (2002). *Journal of Archaeological Science*, 30: 381–6.
- Domínguez-Rodrigo, M. and Pickering, T. R. 2003. Early hominids, hunting and scavenging: a summary of the discussion. *Evolutionary Anthropology*, 12: 275–82.
- Domínguez-Rodrigo, M. and Barba, R. 2005. A study of cut marks on small-sized carcasses and its application to the study of cut-marked bones from small mammals at the FLK Zinj site. *Journal of Taphonomy*, 2–3: 111–24.
- Domínguez-Rodrigo, M., Egeland, C. P. and Barba, R. 2007. *Deconstructing Olduvai*. New York: Springer.
- Fernández-López, S. 2006. Taphonomic alteration and evolutionary taphonomy. *Journal of Taphonomy*, 4(3): 111–42.
- Gifford, D. 1977. Observations of modern human settlements as an aid to archaeological interpretation. Doctoral dissertation, Department of Anthropology, University of California, Berkeley.
- Gifford, D. 1981. Taphonomy and paleoecology: a critical review of archaeology's sister disciplines. In *Advances in Archaeological Method and Theory*, Vol. 4 (ed. M. Schiffer). New York: Academic Press, pp. 365–438.
- Gifford-Gonzalez, D. P. 1989. Ethnographic analogues for interpreting modified bones: some cases from East Africa. In *Bone Modification* (eds R. Bonnichsen and M. H. Sorg). Center for the Study of the First Americans, University of Maine, pp. 179–246.
- Gifford-Gonzalez, D. P. 1991. Bones are not enough: analogues, knowledge, and interpretive strategies in zooarchaeology. *Journal of Anthropological Archaeology*, 10: 215–54.
- Gould, R. A. 1980. *Living Archaeology*. Cambridge: Cambridge University Press.
- Gould, S. J. 1965. Is uniformitarianism necessary? *American Journal of Science*, 236: 223–8.
- Gould, S. J. 1979. Agazziz's marginalia in Lyell's Principles, or the perils of uniformity and the ambiguity of heroes. *Studies in the History of Biology*, 3: 119–38.
- Lakatos, I. 1978. *The Methodology of Scientific Research Programmes*. Cambridge: Cambridge University Press.
- Lupo, K. D. and O'Connell, J. F. 2002. Cut and tooth mark distributions on large animal bones: ethnoarchaeological data from the Hadza and their implications for current ideas about early human carnivory. *Journal of Archaeological Science*, 29: 85–109.
- Lyman, R. L. 1987. Archaeofaunas and butchery studies: a taphonomic perspective. *Advances in Archaeological Method and Theory*, 10: 249–337.
- Lyman, R. L. 2004. The concept of equifinality in taphonomy. *Journal of Taphonomy*, 2(1): 15–26.
- Lyman, R. L. 2005. Analyzing cut marks: lessons from artiodactyle remains in the Northwestern United States. *Journal of Archaeological Science*, 32: 1722–32.
- Marean, C. W., Spencer, L. M., Blumenshine, R. J. and Capaldo, S. 1992. Captive hyaena bone choice and destruction, the Schlepp effect and Olduvai archaeofaunas. *Journal of Archaeological Science*, 19: 101–21.
- Marean, C. W., Domínguez-Rodrigo, M. and Pickering, T. R. 2004. Skeletal element equifinality in zooarchaeology begins with method: the evolution and status of the 'shaft critique'. *Journal of Taphonomy*, 2: 69–98.

- Monahan, C. 1996. Variability in the foraging behavior of early Homo: a taphonomic perspective from Bed II, Olduvai Gorge, Tanzania. Doctoral dissertation, Department of Anthropology, University of Wisconsin, Madison.
- Nairn, A. E. M. 1965. Uniformitarianism and environment. *Palaeogeography, Palaeoclimatology and Palaeoecology*, 1: 5–11.
- Nilssen, P. J. 2000. An actualistic butchery study in South Africa and its implications for reconstructing hominid strategies of carcass acquisition and butchery in the Upper Pleistocene and Plio-Pleistocene. Doctoral dissertation, University of Cape Town, South Africa.
- Pickering, T. R., Marean, C. W. and Domínguez-Rodrigo, M. 2003. Importance of limb bone shaft fragments in zooarchaeology: a response to 'On *in situ* attrition and vertebrate body part profiles' (2002) by M. C. Stiner. *Journal of Archaeological Science*, 30: 1469–82.
- Pobiner, B. 2007. Hominin-carnivore interactions: evidence from modern carnivore bone modification and early Pleistocene archaeofaunas (Koobi Fora, Kenya; Olduvai Gorge, Tanzania). Doctoral dissertation, Department of Anthropology, Rutgers University, New Brunswick.
- Pobiner, B. and Braun, D. 2005. Strengthening the inferential link between cut mark frequency data and Oldowan hominid behavior: results from modern butchery experiments. *Journal of Taphonomy*, 3(2–3): 107–20.
- Popper, K. 1956. *Realism and the Aim of Science*. London: Hutchinson.
- Popper, K. 1972. *Objective Knowledge*. Oxford: Clarendon Press.
- Richter, R. 1928. Aktuopalaontologie und Palaobiologie, eine Abgrenzung. *Senckenbergiana*, 19.
- Selvaggio, M. M. 1994. Identifying the timing and sequence of hominid and carnivore involvement with Plio-Pleistocene bone assemblages from carnivore tooth marks and stone-tool butchery marks on bone surfaces. Doctoral dissertation, Rutgers University.
- Shea, J. H. 1982. Twelve fallacies of uniformitarianism. *Geology*, 10: 455–60.
- Simpson, G. G. 1970. Uniformitarianism: an inquiry into principle, theory, and method in geohistory and biogeography. In *Essays in Evolution and Genetics in Honor of Theodosius Dobzhansky* (eds M. K. Hecht and W. C. Steere). New York: Appleton, pp. 43–96.
- Tappen, M. 1992. Taphonomy of a central African savanna: natural bone deposition in Parc National des Virunga, Zaire. Doctoral thesis, Department of Anthropology, Harvard University, Cambridge, MA.
- Wylie, A. 1982. An analogy by any other name is just as analogical: a commentary on the Gould-Watson dialogue. *Journal of Anthropological Archaeology*, 1: 382–401.
- Wylie, A. 1988. Simple analogy and the role of relevant assumptions: implications of archaeological practice. *International Studies in the Philosophy of Science*, 2: 134–50.
- Wylie, A. 1989. The interpretive dilemma. In *Critical Traditions in Contemporary Archaeology* (eds V. Pinsky and A. Wylie). Cambridge: Cambridge University Press, pp. 18–27.

Manuel Domínguez-Rodrigo is a tenured professor at the Department of Prehistory, Complutense University (Madrid, Spain). He has conducted actualistic neo-taphonomic research in Kenya and Tanzania and archaeological research at Peninj and Eyasi (Tanzania), Gona (Ethiopia) and Swartkrans (South Africa). He is currently co-director of research at Olduvai Gorge under TOPPP (The Olduvai Paleoanthropological and Paleoenvironmental Project). He is also co-editor of the *Journal of Taphonomy*.

Copyright of *World Archaeology* is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.