
Volume 37 Number 2 February 2010 ISSN 0305-4403
ELSEVIER
Published in association with the
Society for Archaeological Sciences
Journal of
Archaeological
SCIENCE

Journal of Archaeological Science 37 (2010) 225–242

Contents lists available at ScienceDirect

Journal of Archaeological Science

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

Review

Assessing the reliability of criteria used to identify mandibles and mandibular teeth in sheep, *Ovis*, and goats, *Capra*

Melinda A. Zeder [a,*], Suzanne E. Pilaar [b]

[a] *Archaeobiology Program, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560-0001, United States*
[b] *Department of Archaeology, University of Cambridge, Downing Street, Cambridge CB2 3D*

A R T I C L E I N F O

Article history:
Received 9 October 2008
Received in revised form
26 September 2009
Accepted 2 October 2009

Keywords:
Ovicaprids
Dental taxonomic criteria
harvest profiles

A B S T R A C T

Criteria developed to distinguish between the mandibles and mandibular teeth of sheep and goats are evaluated using modern specimens from the collections of the National Museum of Natural History and the Field Museum of Natural History. Certain teeth prove to be quite reliable in this regard (dP3, P3, P4). Other dental elements, however, are not reliable and should not be used in dental based identifications of these closely related species (dP4, M1, and mandibular bones). Overall, the identification of sheep dentition using these criteria is more reliable than it is for goats. However, the generally greater likelihood of incorrect identification of goat teeth and mandibles means that specimens identified as sheep are likely to contain high proportion of misidentified goats. Both the proportion of remains that can only be classified as 'Ovis-Capra' and the proportion of incorrectly identified teeth and mandibular bones varies with age, with identifications of younger and older animals less reliable than those of animals in the 1.5–6 year age range. Reliability of identifications increases when more criteria and more elements are used. Identifications based on whole mandibles with complete tooth rows are quite reliable. However, the proportion of specimens likely to be classified as 'Ovis-Capra' also increases in whole mandible identifications, especially in goats. Systematic differences in the proportions of indeterminate and incorrect identifications have a significant impact on species-level dentition-based harvest profiles. Recommendations are made for the use of dentition and long bone based harvest profiles in the study of sheep and goat exploitation strategies.

Published by Elsevier Ltd.

1. Introduction

Developing reliable criteria that can be used to discriminate between the bones and teeth of sheep and goats is one the most enduring and most challenging methodological issues in archaeozoology. These closely related species are found together in archaeozoological assemblages with great geographic and temporal reach. Wild sheep and goat were prominent prey species of Paleolithic hunters across Asia. As domestic animals they are usually well represented in more recent archaeological assemblages wherever these peripatetic species were utilized. Accurate discrimination between the remains of sheep and goat is key to determining the relative importance of each species to the people who exploited them. Since people often exploited sheep and goats in different ways, whether as prey or domesticates, reliable criteria for distinguishing between the bones and teeth of sheep and goat are also an essential step in the construction of species-level harvest profiles capable of detecting these different economic strategies.

Over the years a number of highly effective criteria have been developed to distinguish between the post-cranial elements of sheep and goats (Boessneck et al., 1964; Boessneck, 1970; Hole et al., 1969; Prummel and Frisch, 1986). Combined with increasingly high resolution data on the sequence and timing of long-bone fusion (Silver, 1969; Noddle, 1974; Hatting, 1981; Garcia-Gonzalez, 1983; Moran and O'Connor, 1994; Zeder, 2006a; Bullock and Rackham, 1982), it is now possible to compute robust species-level harvest profiles based on long-bone fusion data (e.g. Zeder, 2008). However, fusion-based harvest profiles cannot project any farther than the age of late fusing bones (i.e. the proximal humerus which fuses at about 4 years of age in sheep and goats). For an animal that may live as long as 10–12 years or more, this is a serious limitation.

Tooth eruption and wear sequences, in contrast, cover the entire lifespan of these animals, from birth to death, and, thus, promise a much more robust profile of the harvest strategies of both hunters and herders. In the early 1970s Payne (1973) published a set of

* Corresponding author.
E-mail address: zederm@si.edu (M.A. Zeder).

0305-4403/$ – see front matter Published by Elsevier Ltd.
doi:10.1016/j.jas.2009.10.002

tooth eruption and wear criteria that could be used in the aging of sheep and goat mandibles. These original criteria have since been subjected to many assessments and enhancements (Deniz and Payne, 1982; Payne, 1987; Jones, 2006; Zeder, 1991, 2006a; Bullock and Rackham, 1982). As a result, archaeozoologists now have a reliable and easily utilized set of criteria for the computation of dentition-based harvest profiles for caprine dentition that covers essentially the entire lifespan of these animals. However, lacking reliable criteria for distinguishing between the teeth of sheep and goats, most dentition-based harvest profiles have, until recently, been computed using combined assemblages of sheep and goat remains – a limitation that restricts the utility of these profiles.

In 1985 Payne addressed this problem by proposing a number of criteria that he believed could be used to distinguish between certain teeth in juvenile sheep and goats (Payne, 1985): the first milk incisor (i_1 or dI_1), the second milk molar (m2 or dP3), the third milk molar (m3 or dP4), and unworn first molar's (M1). This initial attempt was followed by the work of Helmer (2000) who presented criteria for distinguishing between caprine permanent pre-molars – the third pre-molar (P3) and the fourth pre-molar (P4). Helmer's criteria were later incorporated into a more ambitious effort by Halstead et al. (2002) that proposed taxonomic criteria for all adult cheek teeth (P3 through M3) and the mandibular bones of sheep and goats. Balaase and Ambrose (2005) have also contributed to the effort at developing criteria for the discrimination of sheep and goat teeth. Many archaeozoologists have eagerly embraced at least some of these new criteria and are increasingly using them to construct species-level dentition-based harvest profiles used to make fine scale distinctions between different sheep and goat exploitation strategies (e.g. Vigne and Helmer, 2007).

In this paper, we present an independent assessment of the efficacy of these criteria based on a sample of 62 individuals (123 mandibles) of modern sheep and goat from the collections of the Smithsonian Institution's National Museum of Natural History and an additional 22 specimens (44 mandibles) from the Field Museum of Natural History. This assessment was undertaken as part of a larger study of sheep and goat domestication in the eastern Fertile Crescent region of today's Iran and Iraq (Zeder, 2001, 2005, 2006b, 2008). Our initial goal was simply to acquaint ourselves with these criteria on specimens with secure species assignment before we attempted to apply them to archaeological assemblages. However, our examination of these modern specimens raised some questions about the reliability of these criteria and their use in the computation of species-level harvest profiles. We present the results of our assessment and the questions it raises here.

2. Sample and methods

The initial sample examined consisted of 121 mandibles representing 17 goats and 45 sheep from the collections of the National Museum of Natural History (Table 1, Table S1). Subsequently this sample was expanded to include an additional 44 mandibles of younger animals with deciduous teeth, representing nine goats and 13 sheep, from collections of the Field Museum of Natural History. All specimens were associated with skulls and are securely identified at least to genus level as either sheep or goat. The sample includes both males and females and animals that span the full range of age classes, from very young to very old individuals. In addition, 31 of these specimens (15 goats and 16 sheep) are wild animals (*Capra aegargus* and *Ovis orientalis*) from eastern Iran and Pakistan. Eleven specimens were originally identified in the NMNH catalogue as *Ovis europaeus orientalis* or *Ovis musimon*. These specimens are either from the island of Sardinia or represent zoo animals that are most likely the feral ancestors of domestic sheep left behind on Mediterranean islands by early Neolithic colonizers of the Western Mediterranean.

Table 1
Summary of specimens examined in this study by sex, domestic status, and age.

Category	Goat	Sheep	Total
Total	26	58	84
Male	13	30	43
Female	9	21	30
Sex unknown	4	7	11
0–1.5 yrs	12	29	41
2–6 yrs	11	24	35
>6 yrs	3	5	8
Age unknown	0	0	0
Wild	15	16	31
Domestic	10	23	33
Feral	0	11	11
Status unknown	1	8	9

Criteria evaluated are derived from Payne (1985), Helmer (2000), Halstead et al. (2002), and Balaase and Ambrose (2005), with one criterion observed by Pilaar in the course of our study. These criteria are presented in Figs. 1–7. Each tooth on each of the 167 mandibles was scored for each criterion (Table S2). Both the right and left mandibles of each specimen were included in the sample. Though often nearly identical, there were many instances in which dental and mandibular features varied within the right and left sides of individuals, or in which features were obscured on one side but not on the other. Scoring recorded whether the feature observed conformed to criteria consistent with sheep (O), with goat (C), or was not clearly identifiable as either taxon (O/C). Criteria that could not be observed due to absence of a tooth, breakage, or other obscuring factors were recorded as 'x'.

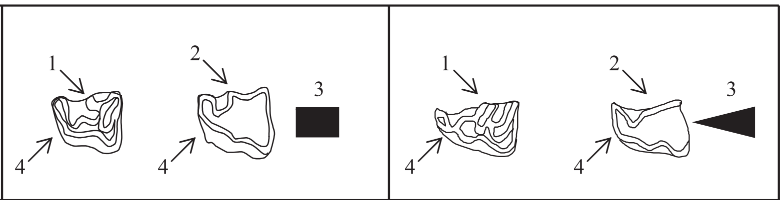
Taxonomic assignment of individual teeth and mandibular bones was made by averaging the results of the individual criterion assessments for each element. If 50% or more of the criteria evaluated were consistent with sheep, for example, the element was scored as 'O'. If there was no majority assignment as 'C', 'O', or 'O/C', then the tooth or mandible was scored as 'O/C'. This method differs from that recommended by Halstead et al. who suggest that whole element identifications be based on a general sense of the combination of observable criteria rather than a simple tally of criteria (Halstead et al., 2002: 550). However, we feel that our scoring system, while more mechanical, provides a more objective assessment of the cumulative value of these criterion, especially since in it less colored by observer bias in making these determinations on mandibles whose taxon is known.

Taxonomic assignment for entire specimens (including all teeth and the mandibular bone) was made in a similar fashion. If the majority of teeth and the mandibular bone were scored as either 'C', 'O', or 'O/C', then the specimen received this assignment. If there was no majority assignment, then the specimen was scored as 'O/C'.

3. Assessment of the entire sample

3.1. Individual criteria

The performance of individual criteria in assigning the correct taxon is presented in Table 2. Individual criteria are listed by element (for each tooth and the mandibular bone), with sample sizes and results presented first for goats and then for sheep. For each taxon, the number of specimens observed is listed, then the percentage of specimens that could not be identified as either sheep or goat but scored as 'O/C' is provided. The next two columns in both the goat and sheep portions of the table present the

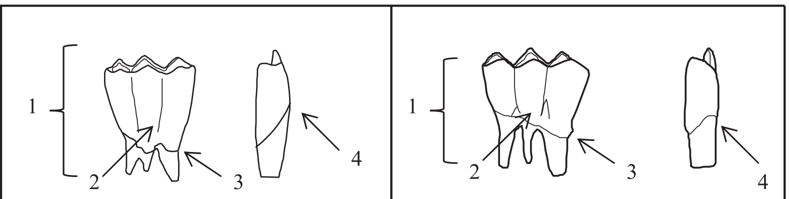


Criterion	Ovis	Capra
1	Unworn & Slightly Worn: Metaconid strong, tipped mesially, linked by short ridge to distal part of tooth. (Payne 1985)	Unworn & Slightly Worn: Metaconid weaker, tipped distally, linked by ridge to more mesial part of tooth. (Payne 1985)
2	Worn: Metaconid projects fairly strongly mesially. (Payne 1985)	Worn: Metaconid essentially disappears. (Payne 1985)
3	Worn & Unworn: Tooth is heavier and squarer. (Payne 1985)	Worn & Unworn: Tooth is more narrowly triangular. (Payne 1985)
4	Worn & Unworn: Bucco-mesial corner of lower part of crown marked by an angle or small ridge. (Payne 1985)	Worn & Unworn: Bucco-mesial corner of lower part of crown marked by a more gradual curve. (Payne 1985)

Fig. 1. Criteria for distinguishing Ovis and Capra dP3.

percentage of correct and incorrect taxon assignments among the remaining specimens. Taking the third criterion evaluated for the P4 (P4:4) as an example, of the 20 goat P4s examined, five (25%) did not definitively display the expected characteristics for either sheep or goat and were scored as indeterminate 'O/C'. Of the remaining 15 specimens, ten (66.7%) displayed the longer, narrower shape predicted for goats by P4:4 and were thus classified correctly as goat. Five specimens (33.3% of the 15 identified specimens) displayed the shorter, broader shape held to be characteristic of sheep and, thus, would in a blind study have been incorrectly identified as sheep. For the same criterion, five (9.8%) of the 51 sheep specimens examined were classified as indeterminate "O/C". Forty of the 46 remaining specimens (87% of the identified specimens) displayed the squarer shape held to be characteristic of sheep, while six specimens (13%) would in a blind study have been wrongly attributed to goat.

The last two columns of Table 2 look at the same data from a somewhat different perspective. Based on the rates of both correct and incorrect assignment, these columns project the proportion of specimens that, for that criterion, would have been positively identified as either or sheep or goat, but that would, in fact, have been the other taxon. Since the number sheep and goat mandibles examined were quite different (there were more than two times as many sheep as goat in our sample), we used normalized percentages of correct and incorrect identifications to compute this figure. Using the same P4:4 criterion once again as an example, based on the rates of correct and incorrect assignment reported for both the goats and the sheep in the normalized samples, 79.7 of the specimens would have been identified as goats (arrived at by adding the percentage of correctly identified goats to the percentage of incorrectly identified sheep). Thus in this normalized population, 66.7 of the animals so



Criterion	Ovis	Capra
1	Less of a tendency for basal swelling. (Payne 1985)	Tendency toward swelling bucco-distally at the base of the crown (Payne 1985)
2	Basal inter-lobar pillars are rarely present. (Payne 1985)	Basal inter-lobar pillars often present, especially between middle and distal lobes. (Payne 1985)
3	Unworn & Slightly Worn: Crown more hypsodont, especially in bucco-distal corner. [Obscured in older specimens.] (Payne 1985)	Unworn & Slightly Worn: Crown less hypsodont, especially in bucco-distal corner. [Obscured in older specimens.] (Payne 1985)
4	Unworn & Slightly Worn: Enamel base rises steeply (usually at 60° angle or more) from the buccal to the lingual side. [Obscured in older specimens.] (Payne 1985)	Unworn & Slightly Worn: Base of enamel rises less steeply (usually at 45° angle or less) from the buccal to the lingual side. [Obscured in older specimens.] (Payne 1985)

Fig. 2. Criteria for distinguishing Ovis and Capra dP4.

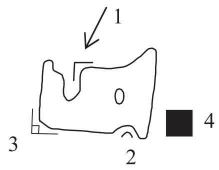
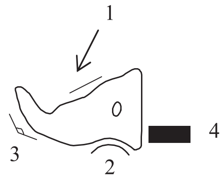
		
Criterion	**Ovis**	**Capra**
1	Vertical ridge in middle of lingual face more strongly developed, usually making the lingual edge of the occlusal face more stepped with clear right angle turn. (Helmer, 2000, Halstead et al., 2002, P3.1)	Vertical ridge in middle of lingual face less strongly developed, so that the lingual edge of the occlusal face inclines more gradually in an antero-posterior direction. (Helmer, 2000, Halstead et al., 2002, P3.1)
2	Vertical ridge on disto-buccal corner more pronounced, so that distal part of buccal edge forms a relatively distinct, deep hollow. (Halstead et al., 2002, P3.2)	Vertical ridge on disto-buccal corner less pronounced, so that distal part of buccal edge forms a more shallow and less distinct hollow. (Halstead et al., 2002, P3.2).
3	Mesial part of buccal face slopes less strongly inwards (lingually) in a postero-anterior direction, so that the mesial face is typically perpendicular to the mandible and the mesio-buccal quarter of the tooth tends toward a right angle. The mesio-buccal "corner" may be rounded or angled in both species. (Helmer, 2000, Halstead et al., 2002, P3.3)	Mesial part of buccal face slopes more strongly inwards (lingually) in a postero-anterior direction, so that the mesial face slopes anteriorly in a bucco-lingual direction and the mesio-buccal quarter of the tooth tends toward a more open angle (approaching a straight line in extreme cases). The mesio-buccal "corner" may be rounded or angled in both species. (Helmer, 2000, Halstead et al., 2002, P3.3).
4	Overall tooth is shorter and broader. (Helmer, 2000Halstead et al., 2002, P3.4).	Overall tooth is longer and more slender. (Helmer, 2000Halstead et al., 2002, P3.4)

Fig. 3. Criteria for distinguishing *Ovis* and *Capra* P3.

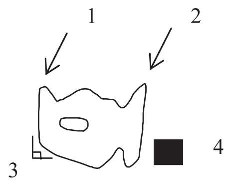
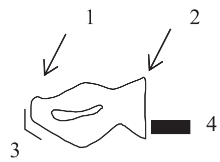
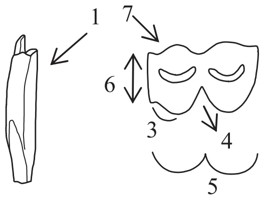
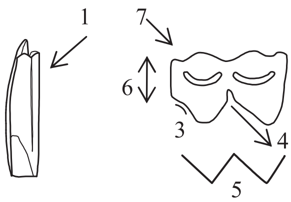
		
Criterion	**Ovis**	**Capra**
1	Mesio-lingual corner typically marked by a vertical rib that projects lingually. (Helmer, 2000, Halstead et al., 2002, P4.1)	Mesio-lingual corner usually lacking vertical rib that projects lingually. (Helmer, 2000, Halstead et al., 2002, P4.1)
2	Disto-lingual corner sometimes has a vertical rib that projects lingually. (Halstead et al., 2002, P4.1)	Disto-lingual corner usually lacks vertical rib that projects lingually. (Halstead et al., 2002, P4.1)
3	Mesio-buccal quarter of tooth forms a right angle that is sometimes rounded and sometimes with the "corner" emphasized by a vertical rib. (Helmer, 2000, Halstead et al., 2002, P4.2)	Mesio-buccal quarter of tooth forms an obtuse angle (clearly > 90º). (Helmer, 2000, Halstead et al., 2002, P4.2)
4	Overall tooth is shorter and broader, but shape may be obscured by wear and tooth crowding. (Helmer, 2000, Halstead et al., 2002, P4.3).	Overall tooth is longer and slenderer, but shape may be obscured by wear and tooth crowding. (Helmer, 2000, Halstead et al., 2002, P4.3)

Fig. 4. Criteria for distinguishing *Ovis* and *Capra* P4.





Criterion	**Ovis**	**Capra**
1	M1 only: Unworn: Mesial fold narrows a little below the top of the crown, then widens again. Width of the mesial fold larger than in *Capra*. [Disappears with wear, after Zeder wear stage 14-16] (Payne 1985)	M1 only: Unworn: Mesial fold narrows gradually near the top of the crown. Width of the mesial fold smaller than in *Ovis*. [Disappears with wear, after Zeder wear stage 14-16] (Payne 1985)
2	M1 only: Interlobar pillars rarely present. (Payne 1985). (Halstead et al., 2002, M1&2.4).	M1 only: Interlobar pillars sometimes present (3 of 12 in Payne 1985). Sometimes also seen in M2 and M3, but very rare. (Halstead et al., 2002, M1&2.4)
3	Mesial part of buccal edge (ignoring buccally projecting flange on mesial corner) is often convex. May appear hollow or goat-like in unworn M1 and M2 and in heavily worn M1 and M2 (> Zeder wear stage 17). (Halstead et al., 2002, M1&2.1)	Mesial part of buccal edge (ignoring buccally projecting flange on mesial corner) is often concave or hollow. Sheep may appear hollow or goat-like in unworn M1 and M2 and in heavily worn M1 and M2 (> Zeder wear stage 17). (Halstead et al., 2002, M1&2.1)
4	Buccal edge of disto-buccal cusp typically symmetrical, though may show a slight posterior orientation. (Halstead et al., 2002, M1&2.2).	Buccal edge of disto-buccal cusp often points strongly in posterior direction, although may show a slight or no posterior orientation. In unworn or slightly worn teeth (up to Zeder wear stages 13 or 14), the distal margin may flare in a posterior direction so that the buccal edge of the disto-buccal cusp may appear symmetrical or sheep-like (Halstead et al., 2002, M1&2.2).
5	Buccal edge has a rounded "arcaded" appearance. Rounded profile diagnostic of sheep (Halstead et al., 2002, M1&2.3).	Buccal edge has a pointed "triangular" appearance. Flat and pointed profile diagnostic of goats (Halstead et al., 2002, M1&2.3).
6	Flange on mesial face tends to be broad. [Feature is heavily influenced by occlusal wear. Broad flange is suggestive of sheep in lightly worn M1&M2 (up to stage 16) but could be either sheep or goat in M1&M2 with wear ≥ Zeder wear stage 17.] Curvature of mesial face less pronounced. (Payne 1985 for M1, Balaase and Ambrose, 2005 for M2 and M3, Halstead et al., 2002 for M3 applied here to M1&2)	Flange on mesial face tends to be narrow. [Feature is heavily influenced by occlusal wear. Narrow flange is suggestive of goat in medium to heavily worn M1&2 (≥ Zeder wear stage 17), but less securely indicative of goat in lightly worn M1&2.] Curvature of mesial face more pronounced. (Payne 1985 for M1, Balaase and Ambrose, 2005 for M2 and M3, Halstead et al., 2002 for M3 applied here to M1&2)
7	Neck of mesial flang narrows before curving out to form flange, creating a "U" between flange and distal part of tooth on both lingual and buccal sides. (Observed by SP)	Neck of mesial flange is at the same level as the flange, creating a somewhat straight line on the lingual and buccal sides of the flange. (Observed by SP)

Fig. 5. Criteria for distinguishing *Ovis* and *Capra* M1 and M2.

identified would have actually been goats, while 13 would have actually been misidentified sheep. Of the 120.3 animals that would have been identified as sheep using this characteristic, 33.3 would have been misidentified goats. So the rate of misidentified specimens among the specimens identified as goats using this criterion would be 27.7% (33.3/120.3), while 16.4% (13/79.7) of the specimens identified as sheep would have been misidentified goats.

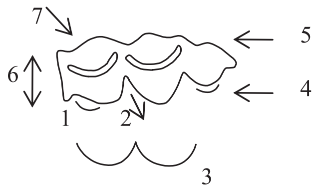
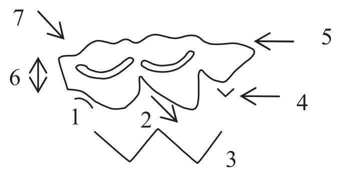
		
Criterion	**Ovis**	**Capra**
1	Mesial part of buccal edge of mesio-buccal cusp (ignoring the buccally projecting phlange on the mesio-buccal corner) is typically convex. Mesial part of buccal edge may be somewhat concave or goat-like in heavily worn M3 (> Zeder wear stage 17). (Halstead et al., 2002, M3.1)	Mesial part of buccal edge of mesio-buccal cusp (ignoring the buccally projecting phlange on the mesio-buccal corner) is typically concave. Sheep may appear to be somewhat concave or goat-like in this feature in heavily worn M3 (> Zeder wear stage 17). (Halstead et al., 2002, M3.1)
2	Buccal edge of the centro-buccal cusp is relatively symmetrical, though may show a slight posterior orientation. (Halstead et al., 2002, M3.2)	Buccal edge of the centro-buccal cusp often points in a posterior direction, although it may show only a slight or no posterior orientation. Distal margin of the centro-buccal cusp of an unworn or lightly worn tooth (up to Zeder wear stage 13) flares posteriorally, making the buccal edge of the cusp appear symmetrical or sheep-like. (Halstead et al., 2002, M3.2)
3	Buccal edge of the mesial and central parts tend to have a rounded “arcaded” appearance. Rounded profile more diagnostic of sheep but less in M3 than in M1 or M2. (Halstead et al., 2002, M3.3)	Buccal edge of the mesial and central parts tend to have a pointed “triangular” appearance. Flat and pointed profile diagnostic of goats. (Halstead et al., 2002, M3.3)
4	Buccal edge of the distal cusp typically rounded. (Halstead et al., 2002, M3.4)	Buccal edge of the distal cusp often more or less pointed, though may be rounded. (Halstead et al., 2002, M3.4)
5	Distal margin of distal cusp often has a buccally defined flute. Presence of this feature is “strongly suggestive” of sheep. [Flute that can only be seen lingually is of no diagnostic value. Buccally defined flute may be clearer in the buccal than in the occlusal view.] (Halstead et al., 2002, M3.5)	Distal margin of distal cusp rarely has a buccally defined flute. Absence of this feature is “suggestive” of goat. [Flute that can only be seen lingually is of no diagnostic value. Buccally defined flute may be clearer in the buccal than in the occlusal view.] (Halstead et al., 2002, M3.5)
6	Flange on mesial face tends to be broad. [Feature heavily influenced by occlusal wear. Broad flange suggestive of sheep in lightly worn M3 (up to stage 16) but could be either sheep or goat in M3 with wear ≥ stage 17.] Curvature of mesial face less pronounced. (Balaase & Ambrose, 2005, Halstead et al., 2002, M3.6)	Flange on mesial face tends to be narrow. [Feature heavily influenced by occlusal wear. Narrow flange suggestive of goat in medium to heavily worn M3 (≥ stage 17), but less securely indicative of goat in lightly worn M3.] Curvature of mesial face more pronounced. (Balaase &Ambrose, 2005, Halstead et al., 2002, M3.6)
7	Neck of mesial flange narrows before curving out to form flange, creating a “U” between flange and distal part of tooth on both lingual and buccal sides. (SP).	Neck of mesial flange is at the same level as the flange, creating a somewhat straight line on the lingual and buccal sides of the flange. (SP).

Fig. 6. Criteria for distinguishing *Ovis* and *Capra* M3.

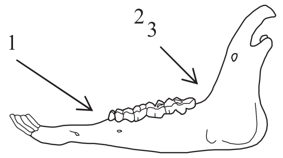
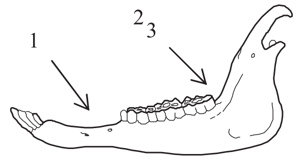
		
Criterion	**Ovis**	**Capra**
1	Foramen often occurs on lateral face of mandible below the P2-4. (Halstead et al., 2002, MD.1)	Foramen usually absent on lateral face of mandible. If it occurs it is typically found anterior to the P2. (Halstead et al., 2002, MD.1)
2	Immediately posterior to M3 on the lateral face of the mandible there is no depression or hollow. A single obvious sinus is often present. [Criterion can only be evaluated after M3 is fully erupted and mandible has achieved full size.] (Halstead et al., 2002, MD.2)	Immediately posterior to M3 on the lateral face of the mandible there is a more or less pronounced hollow. Hollow has no sinus. [Criterion can only be evaluated after M3 is fully erupted and mandible has achieved full size.] (Halstead et al., 2002, MD.2).
3	Same as above. (Halstead et al., 2002, MD.2)	Hollow may have two or more indistinct sinuses. [Criterion can only be evaluated after M3 is fully erupted and mandible has achieved full size.] (Halstead et al., 2002, MD.2)

Fig. 7. Criteria for Distinguishing *Ovis* and *Capra* Mandibles.

Thus the figures presented for sheep and goats in the middle of the table provide a sense of the degree of resolution of each criterion for each species (in the proportion of specimens that cannot be identified to taxon using the criteria) and both the accuracy and error rates of these calls (in the percentage of correct and wrong assignments). The final two columns, on the other hand, combine the separate accuracy and error rates for sheep and goats to project the impact of erroneous identifications on the sample of identified specimens. These final columns, then, give one an appreciation for the degree of error expected in the application of these criteria to an actual sample of unknown specimens, a situation that more accurately reflects potential success, and error, rates of these criteria in an archaeological assemblage.

The dangers of using individual criteria for making these determinations become readily apparent in examining Table 2. Certain criteria (i.e. dP3:1, dP3:2, dP3:4, P3:1, P4:2, P4:3, and M3:1) have high, greater than 80% success rates in both goats and sheep, and subsequent low proportions, under 20%, of projected misidentified specimens among the specimens identified as either goats or sheep on the basis of the criterion. This kind of cross species consistency, however, is uncommon. More often criteria that have high rates of correct assignment for one species do not have similar success rates in the other species. Criterion P4:1, for example, assigned 91.5% of the sheep examined to the correct taxon, while the success rate of this criterion, at 60%, was much lower in goats. Thus despite its high success rate in correctly identifying sheep, the pool of specimens identified as sheep using this criterion would have contained 30.4% misidentified goats.

Average rates for criterion based assessments presented in Table 3 are computed by totaling the percentages in each column in Table 2 and dividing the result by the number of criteria observed. These averages provide an overall sense of the accuracy or error rates of identifications made on the basis of isolated criteria and highlight systematic differences between sheep and goats in the success rates of these criteria in making correct species identification. Single criteria for identifying teeth and mandibular bones are somewhat more likely to be classified as indeterminate in goats than in sheep (10.4% ‘O/C’ in goats versus 8.1% in sheep). When a taxonomic assignment was possible, the accuracy of these identifications in goats is, on average, lower than it is in sheep (71.1% versus 80%). As a result despite the greater degree of accuracy in the identification of sheep teeth using single criteria, the greater potential for misidentification of goats means that specimens identified as sheep are projected to include, on average, 23.1% misidentified goats, while specimens identified as goats are projected to include a slightly lower proportion of misidentified sheep (20.6%).

3.2. Individual elements

As noted by both Payne (1985) and by Halstead et al. (2002), identifications are most reliably made not on the basis of individual criteria, but on the basis of a number of criteria considered taxonomically diagnostic for each element. It is useful, then, to examine the efficacy of identifications made on the basis of a combination of individual criteria for teeth and jaws. Table 4 presents the results of this assessment by element (for each tooth and the mandibular bone (MD)), and Table 5 presents the average of rates of indeterminate, correct, wrong, and misidentified specimens made on the basis of combined trait assessments of individual teeth and mandibles. Somewhat unexpectedly, identifications of teeth and mandibles made on the basis of combined criteria are not always more robust than identifications based on individual criteria.

Success rates for certain teeth are quite high. Both sheep and goats had success rates of 90% or more on the basis of combined traits in two of the eight elements examined (dP3, P4), with two additional elements (P3 and M2) successfully assigned to taxon in more than 80% of the specimens examined. The projected proportion of misidentified specimens for all of these elements is less than 20% in both sheep and goats. Identifications of sheep and goat using these teeth seem to be quite reliable.

Other teeth, however, have much lower success rates. The dP4, in particular, seems a highly unreliable element for distinguishing between sheep and goats. Payne lists four criteria in the dP4 that can be used to discriminate between sheep and goat (Fig. 2). We

Table 2
Results for individual criteria for entire sample.

Tooth	Criteria	Goats				Sheep				% Misidentified	
		#	% O/C	% Corr	% Wrng	#	% O/C	% Corr	% Wrng	Sheep among goats	Goats among sheep
dP3	1	6	16.7	100	0	25	8	100	0	0	0
	2	15	6.7	100	0	32	12.5	85.7	14.3	12.5	0
	3	23	13	95	5	53	39.6	75	25	20.8	6.3
	4	23	0	82.6	17.4	57	14.0	87.8	12.2	12.9	16.5
dP4	1	20	35	38.5	61.5	37	16.2	64.5	35.5	78	48.8
	2*	23	0	21.7	78.3	49	0	81.6	18.4	45.8	48.9
P3	1	21	0	85.7	14.3	51	0	92.2	7.8	8.4	13.4
	2	21	4.8	100	0	51	2	70	30	23.1	0
	3	21	9.5	89.5	10.5	51.	7.8	66.0	34	27.6	13.8
	4	22	18.2	72.2	27.8	51	19.6	70.7	29.3	28.8	28.2
P4	1	23	13	60	40	51	7.8	91.5	8.5	12.4	30.4
	2	21	19.0	88.2	11.8	53	5.7	86	14	13.7	12
	3	23	4.3	95.5	4.5	53	7.5	85.7	14.3	13	5
	4	20	25	66.7	33.3	51	9.8	87	13	16.4	27.7
M1	2*	36	0	5.6	94.4	81	0	100	0	0	48.6
	3	40	22.5	58.1	41.9	82	8.5	86.7	13.3	18.7	32.6
	4*	41	2.4	47.5	52.5	82	3.7	82.3	17.7	27.2	39
	5	45	6.7	78.6	21.4	84	27.4	62.3	37.7	32.4	25.6
	6	28	14.3	83.3	16.7	46	10.9	68.3	31.7	27.6	19.6
	7	23	13	70	30	43	20.9	85.3	14.7	17.4	26
M2	3	30	3.3	72.4	27.6	60	0	96.7	3.3	4.4	22.2
	4*	31	6.5	41.4	58.6	63	0	92.1	7.9	16.1	38.9
	5	31	3.2	83.3	16.7	63	17.5	57.7	42.3	33.7	22.4
	6	25	0	56	44	52	0	84.6	15.4	21.6	34.2
	7	19	15.8	81.3	18.8	51	5.9	75	25	23.5	20
M3	1	22	13.6	84.2	15.8	54	0	96.3	3.7	4.2	14.1
	2*	22	13.6	63.2	36.8	52	0	100	0	0	26.9
	3	22	27.3	75	25	54	3.7	76.9	23.1	23.5	24.5
	4*	17	17.6	21.4	78.6	51	3.9	95.9	4.1	16	45
	5	18	11.1	75	25	52	19.2	42.9	57.1	43.2	36.8
	6*	18	0	55.6	44.4	52	3.8	76	24	30.2	36.9
	7	15	13.3	92.3	7.7	48	4.2	69.6	30.4	24.8	10
MD	1	52	0	88.5	11.5	115	0	61.7	38.3	30.2	15.7
	2	20	15	100	0	56	5.4	77.4	22.6	18.5	0
	3	18	0	83.3	16.7	50	0	70	30	26.5	19.2

*Identified as asymmetrical criteria in Halstead et al. (2002).

evaluated two of these criteria (dP4:1 and dP4:2) on the NMNH and FMNH specimens. The first (dP4:1) concerns the degree of basal swelling, especially on the bucco-distal corner of the tooth, held to be more pronounced in goats. The second (dP4:2) is based on the presence or absence of an inter-lobar pillar on the lingual side of the tooth, said to be often present in goats but generally absent in sheep. These criteria could be evaluated on most of the younger animals where the dP4s were fully erupted or where, in especially young specimens, the tooth could be pulled up in the jaw sufficiently to observed the enamel base of the tooth (Table S2). We did not evaluate the other two criteria proposed by Payne: dP4:3, which involves the degree of hypsodonty (said to be greater in sheep), or dP4:4, which relies on an evaluation of the angle of the enamel as it rises from the buccal to the lingual side of the tooth (said to be steeper in sheep than in goats). Both these criteria require complete removal of the teeth from the mandible, which was not possible with the museum specimens examined here. Moreover, as Payne (1985: 143) notes, these criteria are likely to be

Table 3
Averages by individual criteria.

Sample	Goats				Sheep				% Misidentified	
	#*	% O/C	% Corr	% Wrng	#	% O/C	% Corr	% Wrng	Sheep among goats	Goats among sheep
All	52	10.4	71.7	28.3	115	8.1	80	20	20.6	23.1
All – assy**	52	11.2	78.6	21.4	115	9.4	77.7	22.3	21.8	19.8
0–1.5 yrs	24	7.5	61.9	38.1	57	14.1	84.8	15.2	16.4	24.6
1.5–6 yrs	20	9.8	71	29	48	5.7	81.1	18.9	18.2	22.9
>6 yrs	8	10.1	61.2	35.1	10	4.1	72.1	27.9	27.2	24.9
Wild	30	9.7	69.7	30.3	32	8.8	73.6	26.4	23.5	27.7
Domestic	20	11.8	74.7	25.3	46	8.8	83.4	16.6	18.3	17.6

*# – number of mandibles; ** – All – Assy = entire sample minus “asymmetrical” criteria.

Table 4
Results for individual elements for entire sample.

Tooth	Goats				Sheep				% Misidentified	
	#	% O/C	% Corr	% Inc	#	% O/C	% Corr	% Incr	Sheep among goats	Goats among sheep
dP3	23	13	95	5	57	15.8	91.7	8.3	8.1	5.2
dp4	23	65.2	25	75	49	30.6	79.4	20.6	45.2	48.6
P3	22	4.5	85.7	14.3	53	35.8	91.2	8.8	9.3	13.5
P4	25	36	93.8	6.3	53	15.1	95.6	4.4	4.5	6.1
M1	46	45.7	52	48	81	19	94.1	5.9	10.2	33.8
M2	32	25	83.3	16.7	63	3.2	93.4	6.6	7.3	15.1
M3	22	27.3	75	25	54	5.6	100	0	0	20
MD	52	1.9	88.2	11.8	115	4.3	67.3	32.7	27.1	14.9

obscured in more heavily worn specimens (i.e. animals greater than 6 months of age) limiting their potential utility to a small subset of very young animals. Even in the less worn teeth of younger animals these criteria seem quite subjective and open to observer error.

The dP4 criteria we could evaluate did not perform well, especially the widely used second criterion, the inter-lobar pillar (dP4:2) held to be "often present" in goats but "rarely" present in sheep (Payne, 1985: 143). Of the 23 goat mandibles with deciduous teeth examined only five (21.7%) had evidence of inter-lobar pillars (often fairly weakly expressed) while the majority (18 specimens, 78.3%) had no such pillars (Table 2, Fig. 8). These pillars were present, and sometimes strongly expressed, in nine of the 49 young sheep mandibles examined (18.4%) (Fig. 8), occurring roughly about as frequently as in the young sheep examined and in goats (21.4% of young goats had this feature compared to 18.4% in sheep). These results do not seem to be an artifact of the sample of animals studied. Deciduous P4s bearing inter-lobar pillars, for example, were observed in domestic sheep, wild sheep, and feral sheep. Similarly, goats lacking pillars were not confined to any one sub-group of the sample examined. The reliability of the other dP4 criterion evaluated, the degree of basal swelling (dP4:1), is hampered by both the qualitative nature of this character, which results in a high proportion of specimens classified only as O/C (35% in goats and 16.2% in sheep), and high error rates in taxonomic assignment in both goats (61.5%) and sheep (35.5%). When considered together in making whole element taxon assignments (Table 4), the cumulative impact of these poorly performing criteria is a high proportion of unidentifiable dP4s in both goats and sheep (66.7% and 30.6% respectively) and high error rates in the specimens assigned to taxon, especially in goats (75%). Incorrectly assigned dP4s approach 50% in both sheep and goats.

The M1 was another poor performer (Table 4). Nearly half of the goat M1s examined could only be identified as O/C, and only 52% of the remaining goat M1s were classified as goat based on the combined results for the seven different criteria evaluated here. The large number of incorrect identifications of goat M1s would result in a projected high rate of misidentified specimens among M1s identified as sheep (33.8%).

Moreover, the percentage of specimens that can only be identified as 'O/C' increases when single criteria are combined to make identifications for individual teeth and mandibles. This is especially true for goats where the proportion of 'O/C' identifications increases from an average of 10.4% when criteria are considered singly (Table 3) to 27.3% when combined by element (Table 5). The proportion of indeterminate identifications also increases in sheep, from 8.1% to 16.2%. The disparity between the proportions of indeterminate identifications in goats compared to sheep also grows substantially when criteria are combined by element. Using single criteria there was only a roughly 2% difference between the proportions of indeterminate assignments in sheep and goat; with whole element assignments there is an 11% difference between sheep and goats in the proportion of specimens identifiable only as 'O/C'.

The average percentage of correctly identified goats made on the basis of whole element identifications at 74.8% (Table 5) is not much different from what it is using isolated criteria (71.7%, Table 3). For sheep, however, the average success rate of identifications rises nearly 10% over the average percentage of correct identifications using isolated criteria (from 80% to 89.1%). On average, the percentage of misidentified specimens using whole element identifications is lower than when identifications are based on individual criterion, but still approaches 20% among sheep, and over 10% among specimens identified as goats.

3.3. *Whole mandibles*

The researchers who developed these criteria also predicted that taxonomic assignments made on the basis of multiple teeth within single mandibles are much more reliable than those made using either single criteria or combined criteria for single elements. This prediction is born out in Table 6 that presents the results of the whole specimen assignments made on the entire sample studied here. All of the 23 goat mandibles not classified as 'O/C' were correctly identified. In sheep, only five of the 115 identifiable sheep whole mandibles were incorrectly identified as goats. One could, then, be reasonably sure that identification made on the basis of

Table 5
Averages by individual elements.

Sample	Goats				Sheep				% Misidentified	
	#*	% O/C	% Corr	% Wrng	#	% O/C	% Corr	% Wrng	Sheep among goats	Goats among sheep
All	52	27.3	74.8	25.2	115	16.2	89.1	10.9	13.9	19.7
0–1.5 yrs	24	46.8	55.7	24.3	57	15.8	86.1	13.9	17.2	23.9
1.5–6 yrs	20	22.6	82.5	17.5	48	12.7	91.9	8.1	8.1	14.4
>6 yrs	8	16	69.5	30.5	10	9.4	82.9	17.1	27.8	20.9
Wild	30	30.3	72.4	27.6	32	19.1	84.3	15.7	16.9	22.5
Domestic	20	25.3	87.3	12.7	46	17.1	90.8	9.2	8.4	9.9

*# – number of mandibles.

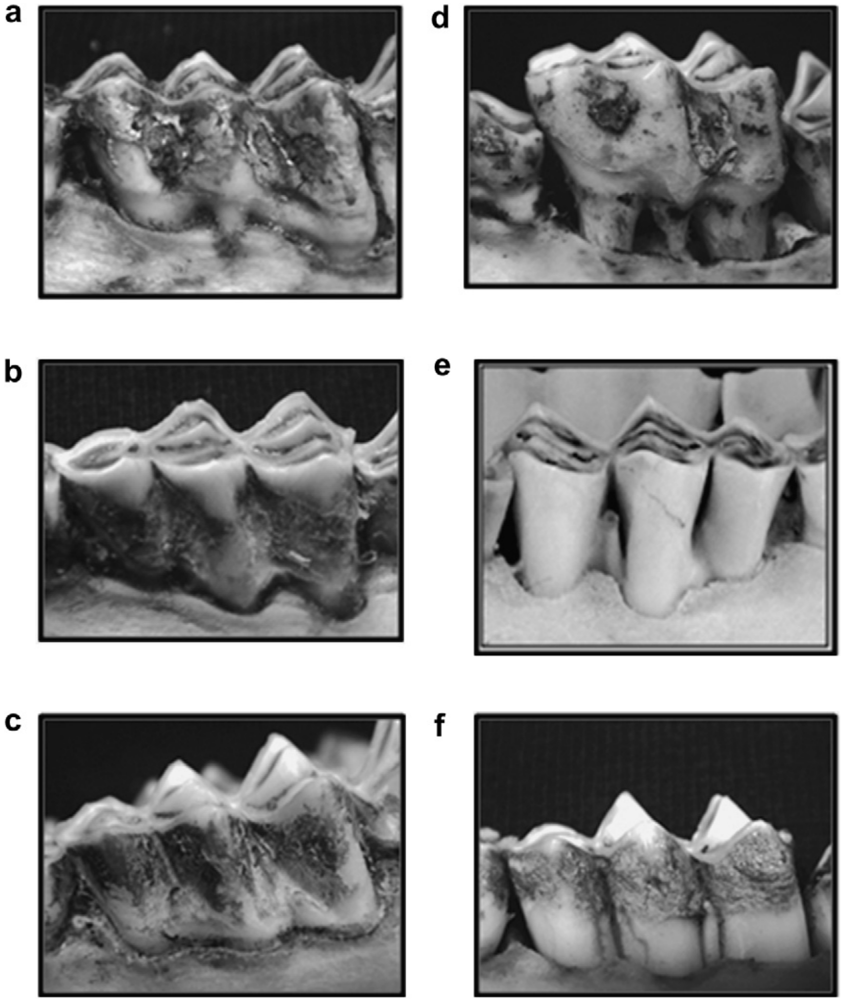


Fig. 8. *Ovis* and *Capra* dP4s. a. *Capra* without inter-lobar pillar but with basal swelling (FMNH 97911), b. *Capra* without inter-lobar pillar or basal swelling (FMNH 180661), c. *Capra* with slight inter-lobar pillar (FMNH 97916), d. *Ovis* with basal swelling (FMNH 58022), e. *Ovis* with strong inter-lobar pillar (NMNH 284889), f. *Ovis* with strong inter-lobar pillar (FMNH 58057).

a whole mandible would be highly accurate. The utility of whole mandible identifications was further demonstrated by the results of a blind study we conducted in which both authors working independently correctly identified all of 10 mandibles selected at random by another member of our laboratory.

Although the majority of the species-level identifications made using whole mandibles were correct, a high proportion of mandibles could not be classified as either sheep or goat but were instead classified as 'O/C' (Table 6). This was especially the case for goats where 55.8% of the mandibles examined could not be assigned to taxon. In contrast, the proportion of whole sheep mandibles identified as 'O/C' was substantially less at 19.1%.

4. Assessment for different age classes

Criteria for distinguishing between the teeth of sheep and goat are highly age dependent. Some of these criteria pertain only to deciduous teeth which are shed within the first two years of life. The utility of other criteria depends on the state of wear of teeth, with some designed specifically for unworn teeth (dP3:1, dP4:1,

Table 6
Results for whole mandibles.

Sample	Goats				Sheep				% Misidentified	
	#	% O/C	% Corr	% Wrng	#	% O/C	% Corr	% Wrng	Sheep among goats	Goats among sheep
All	52	55.8	100	0	115	19.1	94.6	5.4	5.1	0
0–1.5 yrs	24	66.7	100	0	57	29.8	87.5	12.5	11	0
1.5–6 yrs	20	35	100	0	48	6.3	100	0	0	0
>6 yrs	8	75	100	0	10	20	100	0	0	0
Wild	30	66.7	100	0	32	28.1	82.6	17.4	14.8	0
Domestic	20	45	100	0	46	10.9	97.6	2.4	2.4	0

M1:1, and M2:1) and others specific to teeth in a more advanced state of wear (dP3:2–4). We thought it useful, then, to segregate our sample by age and perform the same assessment of the reliability of these criteria for single criteria, for single elements, and for whole specimens. Three general age classes were examined: 0–1.5 year old animals with deciduous teeth, 1.5–6 year old animals with permanent pre-molars and molars in light to moderate wear, and animals greater than 6 years of age whose teeth are in a more advanced state of wear.

Age determinations for the NMNH specimens were made using a system devised by Zeder (1991, 2006a) based on the sequence originally developed by Payne (1973). The numbers of sheep and goat specimens in this sample that fall into these categories can be found in Table 1, with individual specimen age assignments in Table S1. Results of our evaluations of these specimens in these different age brackets were computed in the same way as was done for the entire sample. Tables presenting the results of these assessments for individual criteria in the three age classes can be found in Supplementary Tables S3–S5, with the averaged results for identifications by individual criteria for each age class presented in Table 3. Results of assessments made using combined criteria by element are presented in Table 7. Averages for assessments made on individuals elements are presented in Table 5. The results of whole mandible identifications by age class are presented in Table 6. We will focus primarily on the averaged results for combined criteria and for whole mandibles in this discussion (Tables 5 and 6).

4.1. Animals between 0 and 1.5 years old

Identifications of younger animals are relatively unreliable, especially among goats. Nearly half of the individual teeth and mandibular bones of goats in this age class could not be identified to species and were classified as 'O/C' (Table 5). Of the individual teeth and mandibles of young goats that could be taxonomically assigned, nearly a quarter were incorrectly identified as sheep (Table 5). The poor performance of goat mandibles within this age group is not only attributable problems in the identification of the dP4 discussed earlier. The M1s, M2s and mandibular bones of animals in this age bracket also proved unreliable in making accurate taxonomic identifications. In fact, the dP3 was the only tooth that could be used with any degree of confidence in making accurate taxonomic identifications in younger goats.

Younger sheep fare quite a bit better, with certain elements (dP3, M1, and M2) more reliable in making correct identifications than others (Table 7). Yet, despite the high success rate of criteria in correctly identifying young sheep, the high error rate of identifications of young goats elevates the projected rate of misidentified teeth and mandibles among identified young sheep to nearly 20% (Table 5).

Whole specimen identifications are once again more reliable, with no misidentified specimens among the 23 goats examined, and only five of the 57 sheep whole mandible identifications wrongly attributed to goats (Table 6). However, a very high 65.2% of the sample of 0–1.5 year old goats could only be classified as 'O/C', while a much smaller but still substantial 29.8% of the sheep mandibles in this age class could not be taxonomically assigned to either sheep or goat.

4.2. Animals between 1.5 and 6 years old

Halstead et al. (2002: 550) maintain that mandibles with teeth in a moderate state of wear are more easily identified than mandibles with erupting or highly worn teeth. This conclusion is certainly supported here. Both sheep and goats in this age bracket have the highest percentage of correctly identified specimens, with the teeth and mandibles of 1.5–6 year old goats correctly identified an average of 82.5% of the time, and sheep elements correctly identified 91.9% of the time (Table 5). Among the sheep, the success rate of identifications is greater than 90% in all of the teeth of the animals in this age bracket (Table 7). While once again proving less reliable than sheep, success rates for individual teeth and mandibles in goats are generally quite high. As a result, the projected rates of misidentified specimens in this age bracket are the lowest of any age class. Based on our results it can be predicted that, on average, only 8.1% of specimens identified as goats in this age bracket would be misidentified sheep, while specimens identified as sheep would contain only 14.4% incorrectly identified goats (Table 5).

Whole mandible assignments reflect the greater clarity of distinguishing characteristics among animals in the 1.5–6 year old age bracket (Table 6). Nearly 94% of the sheep mandibles in this 1.5–6 year old age bracket could be identified to species, with only 6.3% were classified as 'O/C'. The percentage of indeterminate mandibles among the goats in this age group, at 35%, was also lower than it was for goats in other age brackets. The disparity between

Table 7
Results for individual elements by age.

Age class	Tooth	Goats				Sheep				% Misidentified	
		#	% O/C	% Corr	% Wrng	#	% O/C	% Corr	% Wrng	Sheep among goats	Goats among sheep
0–1.5 yr	dP3	23	13	95	5	57	15.8	91.7	8.3	8.1	5.2
	dP4	23	65.2	25	75	49	30.6	79.4	20.6	45.2	48.6
	M1	18	55.6	75.0	25.0	37	32.4	100	0	0	20
	M2	4	100	0	0	8	0	100	0	0	–
	MD	24	0	83.3	16.7	57	0	59.6	40.4	32.6	21.8
1.5–6 yr	P3	16	0	87.5	12.5	47	40.4	92.9	7.1	7.5	11.9
	P4	19	42.1	100	0	47	14.9	95.0	5	4.8	0
	M1	20	45	63.6	36.4	39	10.3	94.3	5.7	8.2	27.8
	M2	20	15.0	94.1	5.9	46	2.2	91.1	8.9	8.6	6.1
	M3	14	28.6	60	40	44	4.5	100	0	0	28.6
	MD	20	5	89.5	10.5	48	4.2	78.3	21.7	19.5	11.9
>6 yr	P3	6	16.7	80	20	6	0	83.3	16.7	17.2	19.4
	P4	6	16.7	80	20	6	16.7	100	0	0	16.7
	M1	8	25	0	100	8	0	75	25	100	57.1
	M2	8	12.5	57.1	42.9	10	10	88.9	11.1	16.3	32.5
	M3	8	25	100	0	10	10	100	0	0	0
	MD	8	0	100	0	10	20	50	50	33.3	0

the percentage of goat mandibles in this age bracket that could not be determined to species and the percentage of sheep mandibles classified as ‘O/C’, however, is still quite large.

4.3. Animals greater than 6 years old

Predictably, older animals with greater tooth wear are more difficult to classify. While the average rate of indeterminate identifications among teeth and mandibles is at its lowest rate in both older goats and sheep (Table 5), those teeth and mandibles not classified as ‘O/C’ are quite much more likely to be incorrectly identified than in any other age bracket. As a result, specimens identified as either sheep or goat in this older age bracket are both projected to contain more than 20% misidentified specimens.

Looking at whole mandibles (Table 6), the disparity between sheep and goats in the proportion of identifiable specimens is the largest of any age bracket. Among older caprines there is a 50 percentage point difference between sheep and goat in the proportion of identifiable mandibles. This means that in a population of 100 older animals comprised of 50 goats and 50 sheep, only 15 of the 50 goats mandibles would be identified to species with the rest identified as ‘O/C’. In contrast 40 of the 50 sheep mandibles would be identified to species. Thus rather than the actual 50:50 ratio of old goats to old sheep, the high rate of unidentifiable specimens among the goats in this older age bracket would result in a highly skewed proportional representation of 27% goats to 73% sheep among the mandibles identified to species. As we will see below, differences in the proportion of mandibles that can be identified to species among sheep and goats has important implications for the reliability of species-level age profiles drawn on the basis of whole mandible identifications.

5. Domestic versus wild specimens

We wondered whether the mix of wild, domestic, feral, and zoo animals in our assemblage somehow affected the results of our study. We thought it worthwhile, then, to compare the efficacy of these traits in wild and domestic animals. Our wild sample consists of 62 mandibles from 15 goats and 16 sheep from Iran and Pakistan (Table S1). Zoo specimens and feral animals were not used here. Domestic specimens consist of 66 mandibles from 10 goats and 23 sheep, mostly from the US, but also from domestic animals in Africa, Asia, Oceania, and South America. We hoped that the tighter taxonomic and environmental control of the segregated wild and domestic populations would improve the accuracy our classifications and lower the proportions of specimens that could not be taxonomically assigned. The results of this examination are shown in Supplementary Tables S6 and S7 for individual criteria and in Table 8 for the individual elements. Average success rates for wild and domestic animals are presented in Table 3 for individual criteria, in Table 5 for combined criteria by element, and for whole mandibles in Table 6.

In both species, wild specimens seem more likely to be classified as ‘O/C’ than domestic specimens, whether dealing with individual elements (Table 5) or whole mandibles (Table 6). Moreover, the proportion of incorrectly identified teeth and mandibles is, on average, higher among wild sheep and goats than it is in domestic animals (Table 5). Once again this pattern is especially pronounced in goats. Incorrect identifications of teeth and mandibles are, on average, more than 10% higher in wild goats than in domestic goats (Table 5). As a result the projected proportion of misidentified goat specimens among wild sheep (at 22.5%) is substantially higher than it is among domestic specimens (16.9%) (Table 5). The proportion of whole mandibles that can only be classified as ‘O/C’ is, at 66.7%, is also much higher in the wild goat sample than among domestic goats (28.1%) (Table 6).

It is unclear, however, whether the greater effectiveness of these criteria in correctly identifying the domestic sheep and goats in our sample is attributable to their domestic status. If so, these distinguishing features would represent derived traits specific to domestic animals that are not as strongly expressed in wild ones. This would suggest that the criteria developed for distinguishing between the teeth of sheep and goats are best applied to domestic animals, especially those from periods well after initial domestication. However, if these features are derived characteristics of domestic animals, it is curious that domestic sheep, which have been the object even more intensive breeding than goats, do not show as marked a difference when compared to wild sheep.

6. Caveats and cautionary notes

The detailed discussion above and the extensive documentation of our results in the published tables and in the on-line the supplementary information provide a comprehensive assessment of the utility of dental and mandibular criteria for identifying sheep and goats. The interested reader should have more than enough to go on to ferret out information on the resolution and accuracy of specific criteria at multiple levels of analysis. There are, however,

Table 8
Results for individual elements for wild and domestic specimens.

Age class	Tooth	Goats				Sheep				% Misassigned	
		#	% O/C	% Corr	% Wrng	#	% O/C	% Corr	% Wrng	Sheep among goats	Goats among sheep
Wild	dp3	16	18.8	92.3	7.7	22	9.1	80	20	17.8	8.8
	dp4	16	68.8	20	80	22	27.3	87.5	12.5	38.5	47.8
	P3	13	7.7	75	25	8	37.5	60	40	34.8	29.4
	P4	13	23.1	90	10	8	12.5	100	0	0	9.1
	M1	26	50	53.8	46.2	23	43.5	100	0	0	31.6
	M2	16	37.5	90	10	12	0	91.7	8.3	8.5	9.8
	M3	12	33.3	75	25	10	20	100	0	0	20
	MD	30	3.3	82.8	17.2	32	3.1	54.8	45.2	35.3	23.9
Domestic	dp3	5	0	100	0	20	20	100	0	0	0
	dp4	5	80	100	0	14	42.9	75	25	20	0
	P3	9	0	100	0	23	39.1	92.9	7.1	6.7	0
	P4	12	50.0	100	0	23	26.1	94.1	5.9	5.6	0
	M1	20	40	50	50	32	6.3	93.3	6.7	11.8	34.9
	M2	16	12.5	78.6	21.4	32	0	100	0	0	17.6
	M3	10	20	75	25	22	0	100	0	0	20
	MD	20	0	95	5	46	2.2	71.1	28.9	23.3	6.6

more general patterns that emerge from this study which we feel bear special mention.

6.1. General observations on reliability of dental criteria

Our results suggest that certain teeth are highly unreliable indicators of species and should not be used for this purpose. Among the permanent teeth of older animals, molars are generally less reliable than permanent pre-molars in distinguishing between sheep and goats, with the M1 a particularly unreliable tooth for making taxonomic identifications in goats. In all age brackets, species determinations based on mandibular bone are less reliable than those based on teeth, especially in sheep.

Of the deciduous teeth the dP4 is highly unreliable and should not be used to make taxonomic determinations. The two characteristics evaluated here, the degree of basal swelling and the presence of an inter-lobar pillar failed to distinguish between sheep and goats. In fact, both the purported goat-like characteristics were found in sheep almost as frequently as they were observed in goats, and in neither population were these characteristics very frequently observed. The other two proposed criteria not examined here, the degree of hypsodonty, the angle of the enamel as it rises from the buccal to the lingual side of the tooth hold little promise as alternative criterion for making taxonomic determinations of this tooth. Both criteria are based on highly subjective judgments with plenty of room for inter-observer variation. Moreover, both are obscured by wear, limiting their application to very young animals under six months of age. The dP3, however, performed much better. Indeed, the dP3 seems the only element that can be reliably used for this purpose in younger caprines with deciduous teeth.

Our study also indicates that while these criteria generally perform well in the identification of sheep, they are much less diagnostic in goats. Goat teeth and jaws are much less likely to have features that can be clearly classified as either sheep or goat (resulting in high proportions of indeterminate identifications). When identified to species these identifications are more likely to be incorrect. Thus the higher proportion of misidentified goat teeth and mandibles introduces a potentially significant number of misidentified goats into the sample of specimens identified as sheep. In addition, the small proportion of goat elements that can be identified to species magnifies the impact of even a relatively small number of sheep incorrectly identified as goats.

Our study also underscores the importance of using multiple criteria and multiple elements when making identifications. The proportion of correctly identified specimens increases in both sheep and goats when one moves from individual criteria, to individual teeth and jaws, to whole mandibles. Thus taxon assignments made using whole tooth rows in well preserved mandibles are quite reliable and one can be reasonably sure that identifications made of these specimens are correct.

At the same time, however, the proportion of indeterminate specimens also increases when criteria are combined by element and, even more so, when individual tooth and mandibular bone identifications are combined into whole specimens. The increase in indeterminate assignments is especially marked in goats, with the disparity between sheep and goats in the proportion of specimens classified as ‘O/C’ especially high in whole mandible identifications.

6.2. Impact of variable reliability of dental criteria on harvest profiles

Variability in the proportions of indeterminate and incorrectly identified elements in different age brackets have very serious implications for the reliability of tooth-based harvest profiles based on sheep and goat remains identified using these criteria. The impact of these biases on harvest profiles is best demonstrated through two simulations – one projecting the biases introduced by taxonomic assignments based on individual elements (Table 9 and Fig. 9) and the other projecting the impact of biases inherent in assignment made using whole specimens (Table 10 and Fig. 10).

The first simulation starts with two hypothetical populations of 100 goats and 100 sheep. The goats were exploited following a regime that promotes herd security, while the sheep were exploited for both meat and wool. Both populations were culled in ways that maximized the return of these resources (following Redding, 1981 and Payne, 1973). Culling strategies aimed at maximizing herd security emphasize the slaughter of animals younger than 2 years of age, while strategies that maximize for both meat and wool production in sheep emphasize the slaughter of older animals. Hypothetical harvest profiles for these two populations are presented in the first four columns of Table 9 and graphically shown in the left hand histograms in Fig. 9a and b.

We assume that each of the 200 sheep and goats is represented by a single mandible (100 per species) and that the identification success rates of the teeth and mandibular bones of these specimens will be the same as that as the averages for combined criteria assessments presented in Table 5. First we adjust our samples to accommodate the proportion of indeterminate ‘O/C’ identifications for goats and sheep in different age brackets listed in the fifth and eight columns of Table 9. The number of specimens remaining once these indeterminate specimens are subtracted from the sample is presented in the next columns. Harvest profiles computed using these adjusted samples are presented in columns seven and ten of Table 9 and graphically shown in the middle histograms in Fig. 9a and b.

The high proportion of indeterminate specimens among younger goats substantially reduces the number of identifiable elements the 0–1.5 year age bracket, obscuring the strong emphasis on younger animals seen in the actual age distribution. And while the impact of the smaller proportion of indeterminate specimens among the sheep is less than in goats, the different rates of identifiability among these different age brackets both reduces the proportion of younger animals and inflates the proportion of older animals in the sheep harvest profile.

Table 9
Hypothetical harvest profiles for tooth-based identifications compensating for indeterminate and incorrect identifications.

Age Class	Goat actual		Sheep actual		Deduction of O/C						Adjusted for misidentification					
					Goat			Sheep			Goat			Sheep		
	#	%	#	%	% O/C*	#	%	% O/C*	#	%	% Correct*	#	%	% Correct*	#	%
<1.5 yrs	52	52.0	10	10.0	46.8	27.7	42	15.8	8.4	9.5	55.7	16.6	29.5	86.1	19.5	20
1.5–6 years	30	30.0	50	50.0	22.6	23.2	35.1	12.7	43.7	49.5	82.5	22.9	40.8	91.9	44	44.8
>6 years	18	18.0	40	40.0	16	15.1	22.9	9.4	36.2	41	69.5	16.7	29.7	82.9	34.6	35.2
Total	100		120			66			88.3			56.2			98.1	

* From Table 5.

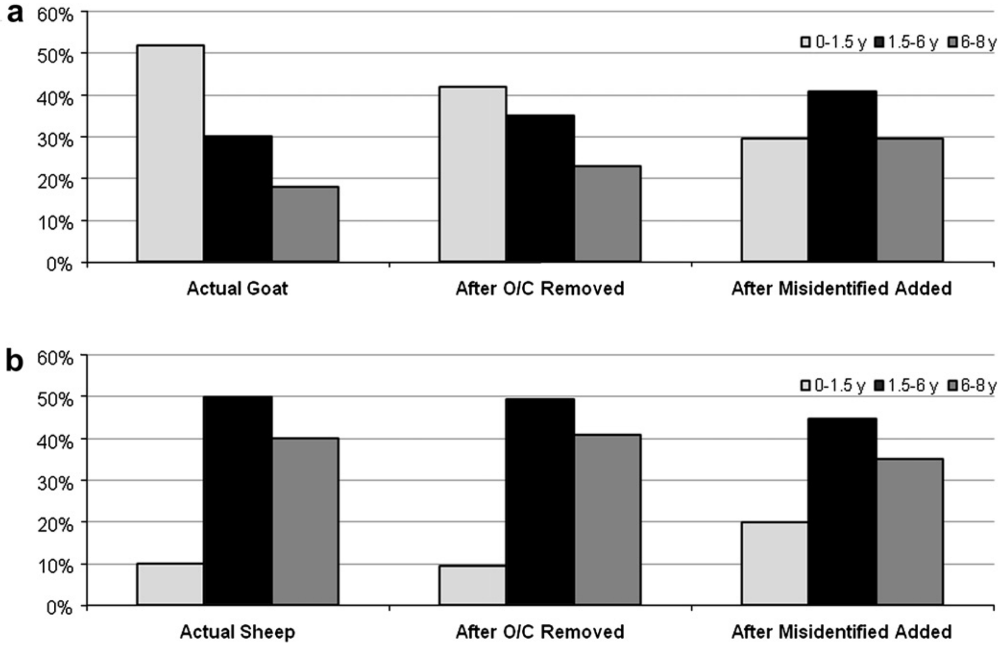


Fig. 9. Simulated Harvest Profiles for *Ovis* and *Capra* based on combined criteria for dental elements. a. *Capra*, b. *Ovis*.

These distributions, however, assume that all of the sheep and goat specimens not classified as indeterminate could be correctly identified to species. We know, however, that not all identifications made on isolated teeth and mandibles are correct, that error rates in goats are generally higher than in sheep, and that error rates vary with age in both species. The percentage of correct identifications for sheep and goats in each bracket are listed in columns 11 and 14 of Table 9 (again derived from Table 5). The next columns present the number of goats and sheep once adjusted for misidentifications (i.e. by adding the misidentified goats to the correctly identified sheep and vice versa).

This later adjustment has a profound impact on the harvest profiles for our populations of sheep and goat. The resultant histograms, displayed in the far right-hand side of Fig. 9a and b, now look quite similar to one another. Once the misidentified elements are added to the mix, the harvest profiles suggest that the sheep and goat in our simulation were harvested (and by extension utilized) in very similar ways, when, in fact, they were managed very differently. The markedly different management regimes of our actual populations are entirely masked by the biases in the proportion of identifiable specimens and in the proportion of correctly identified specimens.

While the whole mandible identifications are much more likely to be correct than those made on the basis of single elements, the higher proportion indeterminate identifications when using whole mandibles also has a transformative impact on harvest profiles. We illustrate the impact of indeterminate attributions in whole specimens in another simulation (Table 10 and Fig. 10). This time we assume that our hypothetical populations of 100 sheep and 100 goats are both managed with an eye toward promoting herd security (Redding, 1981) and both culled in the same proportions in the three age brackets considered here. The harvest profile for these similarly managed animals is graphically represented by the histogram on the left hand side of Fig. 10a and b. The simulation once again assumes that one mandible from each animal was perfectly preserved and analyzed using the criteria evaluated here (100 sheep and 100 goat mandibles in all).

Based on the results of our study, we can project the proportion of whole mandibles for each species and age bracket that would be classified as indeterminate (columns 7 and 10 in Table 10, derived from Table 6). The next columns project the adjusted number of sheep and goat in each age bracket once the indeterminate specimens are removed from the samples, graphically displayed in the right-hand histograms in Fig. 10a and b.

Given the lower proportion of indeterminate mandibles among the sheep, the harvest profile for the sample of identified sheep resembles the actual population, although the higher proportion of identifiable specimens in the 1.5–6 year age bracket, and the misidentifications in the 0–1.5 year age bracket elevates the proportion of animals in older age brackets in the adjusted harvest profile. However, the higher and more variable proportions of indeterminate specimens among the whole goat mandibles in

Table 10
Hypothetical harvest profiles for complete mandibles measuring the impact of indeterminate identifications.

Age class	Goat Actual		Sheep Actual		Deduction of O/C					
					Goat			Sheep		
	#	%	#	%	% O/C*	#	%	% O/C*	#	%
<1.5 yrs	52	52.0	52	52.0	65.2	18.1	43	29.8	36.5	46.2
1.5–6 years	30	30.0	30	30.0	35	19.5	46.3	6.3	28.1	35.6
>6 years	18	18.0	18	18.0	75	4.5	10.7	20	14.4	18.2
Total	100		100			42.1			79	

*From Table 6.

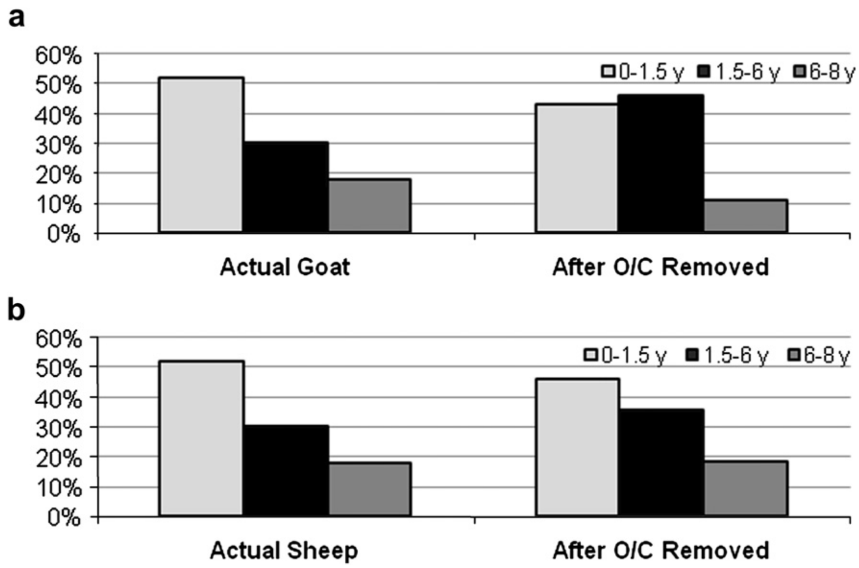


Fig. 10. Simulated harvest profiles for *Ovis* and *Capra* based on whole mandibles.

these different age classes have a profound impact on the goat harvest profile. The adjusted goat harvest profile reduces the original emphasis on 0–1.5 year old animals, over-emphasizes the exploitation of 1.5–6 year old animals, and seriously under represents the exploitation of older animals. The combined impact of these variations in indeterminate identifications in goats produces a harvest profile that might be interpreted reflecting an emphasis on meat maximization (Redding, 1981). In this case, even though all of the identifiable specimens were identified correctly, the bias introduced by differences in the proportions of indeterminate animals results in harvest profiles that suggest very different management regimes for sheep and goat when in actuality sheep and goats were managed in exactly the same way.

6.3. *Comparability between our results and earlier studies*

The results of our assessment of the reliability of dental and mandibular criterion for making taxonomic distinctions between sheep and goat seem at variance with earlier studies that concluded these criteria were highly reliable (Payne, 1985; Helmer, 2000; Halstead et al., 2002; Balaase and Ambrose, 2005). These apparent differences, however, may lie more in the size and nature of the samples examined and in the degree of rigor of the earlier evaluations of their efficacy.

In the Payne (1985) study, for example, methods for distinguishing between deciduous pre-molars were developed using a homogenous sample of 12 sheep and 12 goats comprised almost exclusively of domestic animals purchased from butchers in southern Greece. Payne's comprehensive report on his assessment of these 24 animals makes it possible to assess his results using the methods employed here. Table 11 presents these results for the dP3 and dP4 by individual criteria, by tooth, and for whole specimens based on the cumulative results for these two deciduous teeth. As might be expected, the error rate in the very homogenous sample used to develop these criteria is in almost all cases much lower than in our larger, more heterogeneous sample. However, Payne's assessment echoes ours in several important respects. First, as in our sample, the proportion of indeterminate O/C identifications increases in each different level of assessment (from individual criterion, to whole teeth, to whole specimens). Moreover, with the exception of the dP4:1 criterion, goats always have more indeterminate specimens than sheep, and the disparity between the proportion of indeterminate specimens in the sheep and goat samples increases at each different level of assessment. Indeterminate O/C identifications comprise 66.7% of the goat whole specimen determinations (close to our finding of 65.2% among 0–1.5 year old animals) and only 16.7% of the sheep in Payne's sample.

Helmer's (2000) presentation of criteria used to distinguish between P3 andP4 teeth of sheep and goats was based on a sample of 31 domestic sheep and goats from southern France (with no record of the proportions of sheep and goat in the sample, nor any

Table 11
Payne (1985) results for dP3 and dP4 by criterion, tooth, and whole specimens.

Sample	Goats				Sheep				% Misidentified	
	#	% * O/C	% Corr	% Wrng	#	% O/C	% Corr	% Wrng	Sheep among goats	Goats among sheep
dP3:1&2	12	50	100	0	12	16.7	100	0	0	0
dP3:4	11	36.4	100	0	12	33.3	87.5	12.5	11.1	0
dP4.1	9	11.1	100	0	12	25	100	0	0	0
dP4:2	12	0	91.7	8.3	13	0	84.6	15.4	14.4	9
dP3	12	83.3	100	0	12	58.3	100	0	0	0
dP4	11	54.5	100	0	12	33.3	100	0	0	0
Specimen	12	66.7	100	0	12	16.7	100	0	0	0

*– Includes specimens classified by Payne as indeterminate and as "less typical".

indication of their age or sex), three European ibex, and six feral mouflon from Corsica. He does not, like Payne, present detailed statistics on the results of his study, but he does report that error rates for pre-molars in-wear of about 10% and about 15% in unworn or lightly worn teeth. These error rates are quite similar to (though actually a bit worse) than those in our sample of 1.5–6 year old animals (Table 7). Helmer does not report any statistics on rates of indeterminate O/C determinations.

Balaase and Ambrose's study was based on a control sample of ten goats and six sheep obtained from Massai herders in Kenya. Their metrical analysis of two traits for the identification of the M2 and M3 cannot be compared with ours which is based on observational, not metric data. However, their conclusion that these criteria are most reliable when teeth are in a moderate state of wear does agree with our own results. The author's more impressionistic assessment of the reliability of criteria developed by Helmer (2000) and Halstead et al. (2002) also agrees in most respects with our results. In general, these authors find that criteria used to make taxonomic identification in permanent pre-molars are quite reliable, while those proposed for the identification of the M1 and M2 are problematic. Unlike our study, Balaase and Ambrose found criteria presented in Halstead et al. for the identification of caprine M3s to be "highly diagnostic".

The largest assessment of this kind is presented in the seminal Halstead et al. (2002) study, which developed criteria for taxonomic identification of caprine teeth and mandibles using a sample of 43 adult sheep and 41 goats in the modern reference collection at the University of Sheffield. In addition to this core sample of 84 animals, they also examined the mandibles of 20 younger sheep and 28 goats from the same collections to further assess the use of the M1 and the M2 in species identification. These criteria were evaluated using a sample of 71 sheep (31 adult and 40 young) and 131 goats (107 adult and 24 young) drawn from various museum collections in the United Kingdom and France. The sample included domestic breeds of sheep and goats from Great Britain and continental Europe, as well as domestic goats from Africa and South Asia. Their sample also included both feral and wild sheep and goats. There is, however, no breakdown of the number of animals representing each region, the ages of the animals examined (other than young and adult), or the sexes examined.

Results of their assessment are reported in a summary table (Halstead et al., 2002: Table 2) which presents the number of specimens identified as either sheep or goat and the percentage of correct identifications by individual criteria for the adult pre-molars, the molars, and the mandibular bone. This table also provides the rates of indeterminate identifications, but since indeterminate identifications are reported for the combined sample of sheep and goats, it is impossible to look for differences in the proportion of O/C identifications among sheep and goats. Nor is it possible to evaluate the efficacy of these criteria in determining the taxon of individual teeth and mandibular bones, of whole mandibles, or of subsets of the sample by age or domestic status as was done here. Moreover, the Halstead et al. paper presents no data that would allow one to assess the utility of Payne's criteria for the taxonomic identification of deciduous teeth, though they do report that these criteria are "thoroughly vindicated" in their examination.

One can, however, compare the proportions of correct and wrongly attributed specimens for individual criterion in the Halstead et al. sample with the results of this study. Table 12 presents the Halstead et al. results for individual criteria using the same methods employed here to compute the overall percentage of indeterminate taxon assignments, the percentage of correct and wrong attributions by criterion, and the expected proportions of misidentified animals among the samples specimens assigned to taxon using these criteria. When this is done, some interesting parallels, and differences, can be drawn between the two studies.

Table 12
Results for individual criteria from Halstead et al. (2002).

Tooth	Criteria Z&P	Criteria H et al.	% O/C	Goats			Sheep			% Misidentified	
				#	% Corr	% Wrng	#	% Corr	% Wrng	Sheep among goats	Goats among sheep
P3	1	1	3.8	148	93.9	61	96	85.4	14.6	13.4	6.6
	2	2	5.4	141	94.3	5.7	41	95.2	4.8	4.8	5.6
	3	3	3	137	96.4	3.6	51	92.2	7.8	7.5	3.8
	4	4	12.7	119	94.1	5.9	46	82.6	17.4	15.6	6.6
P4	1	1	1.9	150	96.7	3.3	102	87.3	12.7	11.6	3.7
	3	2	2.4	140	97.9	2.1	60	96.7	3.3	3.3	2.2
	4	3	13.2	121	95.0	5.0	52	86.5	13.5	12.4	5.4
M1	2*	4	.3	30	9.3	6.7	262	43.1	56.9	37.9	13.4
	3	1	12.7	121	93.4	6.6	83	92.8	7.2	7.2	6.7
	4*	2	3.9	121	91.7	8.3	166	60.2	39.8	30.2	12.1
	5	3	7.9	159	98.6	4.4	118	87.3	12.7	11.7	4.8
M2	3	1	5.2	152	95.4	4.6	107	92.5	7.5	7.3	4.7
	4*	2	5.8	99	89.9	10.1	162	60.5	39.5	30.5	14.3
	5	3	12.4	126	95.2	4.8	110	91.8	8.2	7.9	4.9
M3	1	1	5.4	118	96.6	3.4	82	81.7	18.3	15.9	4
	2*	2	4.7	95	93.7	6.3	97	68	32	25.4	8.5
	3	3	13.2	81	100	0	91	75.8	24.2	19.5	0
	4*	4	5.0	74	90.5	9.5	107	51.4	48.6	34.9	15.5
	5	5	2.1	120	87.5	12.5	59	83.1	16.9	16.2	13.1
	6*	6	6.7	103	87.4	12.6	92	64.1	35.9	29.1	16.4
MD	1	1	3.2	222	83.8	16.2	98	94.9	5.1	5.7	14.6
	2	2	5.7	107	100	0	78	85.9	14.1	12.4	0
All			6.2		93.7	6.3		80	20	16.4	7.6
All – Assy			6.9		94.7	5.3		88.2	11.8	10.8	5.4

*– Identified as asymmetrical criteria in Halstead et al. (2002).

As in our study, the Halstead et al. data show that criteria used to identify permanent pre-molars are appreciably more effective than those used to identify molars. Using pre-molar criteria misattributed specimens are always under 15% of the specimens identified as either sheep or goat. Using molar criteria, in contrast, misattributed specimens in their sample increase, with particularly high numbers of incorrectly identified molars among sheep. Halstead et al.'s results differ from ours in the much lower percentage of indeterminate specimens (which varies from about 2–13% of the combined sample of sheep and goats), and in the much higher success rates in the attributions of the taxonomic assignments made using the criteria they evaluated. In our study, of the 29 permanent pre-molar and molar criteria we evaluated only 4 criteria had over 80% correct attributions in both sheep and goats. In contrast, 15 of the 22 criteria evaluated in Halstead et al.'s earlier study met this standard. Moreover, while a mislabeling in the column headings in Halstead et al. Table 2 makes it difficult to interpret their data, it would also appear to that the sheep in their sample are more likely to be incorrectly identified than goats. We found the opposite to be true.

Halstead et al. attribute the lower success rates among sheep in their sample, at least in part, to a number of "asymmetrical" criteria – criteria which, when present, are characteristic of one taxon, but that might be absent in both sheep or goat (Halstead et al., 2002: 549). Success rates for sheep in their sample approaches that found for goats when these criteria are removed from the mix (Table 12). Higher error rates were not restricted to such asymmetrical criteria in our sample. In fact, in many such instances purportedly "goat-like" criteria were found in less than half of the goats examined (i.e. M1:2, M1:4, M2:4, and M3:4) or were equally likely to occur in both sheep and goats (e.g. dP4:2). Averages for indeterminate and incorrect assignments did not improve when these criteria were excluded from consideration in our sample (Table 3); in fact success rates among sheep actually declined when this was done.

It is impossible to identify the precise reason why our two studies differ in the proportions of indeterminate attributions, in overall success rates of the various criteria we evaluated, and in the differences between sheep and goats in the reliability of dental criteria. More data on the nature of the Halstead et al. sample is needed if we are to isolate the exact reasons for the differences between our two studies (i.e. a breakdown of domestic and wild animals in the sample and a report of results by age of animal). Their decision to include identifications of "possible" goat and sheep among the positively identified specimens, rather than to declare these as indeterminate specimens as we were more likely to do, probably lowered the proportion of O/C identifications in their study. Moreover, the fact that these authors both developed and tested these criteria might have increased the accuracy of their application. We, on the other hand, were not as familiar with these criteria as those who first identified them, but were trying to apply them solely on the basis of the descriptions and the figures presented in the Halstead et al. paper. But if taxonomic criteria are to be of any general utility, then other analysts should be able to replicate the results of those who developed them in the first place using such published descriptions. This is, after all, what the majority of archaeozoologists attempting to apply these criteria are doing.

Finally, it is possible that the higher success rates reported by Halstead et al. are attributable to the homogeneity of their sample (i.e. if the sample was comprised primarily of domestic breeds from Western Europe). The accuracy of their determinations might also have been increased if the sample were comprised of adult animals with moderate, but not heavy, states of tooth wear. In fact, the averages of misattributed 1.5–6 year old specimens in our sample (Table 3) are similar to those in the Halstead et al. sample (Table 12), although reversed for sheep and goats.

We submit, however, that if these criteria are to be generally useful in discriminating between sheep and goat across the broad temporal and geographic range of these animals (as is currently being done), then their efficacy needs to be assessed using a heterogeneous sample like ours that includes multiple breeds of domestic animals, wild and feral animals, as well as a variety animals in different age brackets. If on the other hand, these criteria are only effective among certain domesticates from certain parts of the world or in animals that fall into certain age brackets, then this subset of animals needs to be identified and the application of these criteria should be restricted to archaeological sheep and goat remains that fit these parameters. It is often argued that the experienced archaeozoologist knows that caprine assemblages from different geographic areas require different suites of identification criteria, and perhaps that will be the case here. But unless one knows the parameters of breed, taxon, geographic region, and age of animal where such criteria can be reliably used, no amount of experience will allow the analyst to determine whether the taxonomic identifications of unknown archaeological remains made using these criteria are correct or not.

7. Conclusions

Our assessment of criteria used to distinguish between the teeth and mandibles of sheep and goat raises some serious concerns. At the very least, this study suggests that archaeozoologists need to be much more cautious in their use of these criteria in the study of caprine faunal assemblages. Application of these criteria to archaeological assemblages of wild caprines or early domesticates would, based on our assessment, seem particularly risky. Independent assessments of the reliability of these criteria following protocols similar to those adopted here (especially assessments based on Western European domestic caprine breeds) might help isolate a subset of caprines where these criteria might be more effective.

Most importantly, our results indicate that the lack of resolution and accuracy in the application of these criteria can introduce significant biases into species-level harvest profiles based on dental eruption and wear patterns, especially those that focus on the utilization of younger animals. Given the uneven reliability of dental taxonomic assignments observed in our study and the biases they introduce to harvest profiles, any species-level harvest profile based on assemblages identified using these criteria is likely to have been shaped primarily by methodological bias and cannot be considered an accurate reflection of the culling practices of ancient herders.

We conclude that harvest profiles based on dental eruption and wear are only truly reliable in assemblages dominated by one species (as determined on the basis of long-bone taxonomic criteria), and where the entire assemblage of ageable caprine teeth are included in the analysis. Dentition-based harvest profiles are of very limited utility in assemblages where there is a more balanced representation of sheep and goats, especially if these two species were utilized in different ways. As we have demonstrated, using the full-suite of dental taxonomic criteria to construct species-level harvest profiles for sheep and goats can result in profiles that reflect biases in the resolution and accuracy of dental taxonomic criteria rather than ancient harvest practices. Although there are some reasonably reliable dental taxonomic criteria, limiting analysis to this small number of markers would further reduce an already restricted sample of ageable teeth, subjecting any resultant harvest profile to small sample biases. In assemblages with a mix of sheep and goats, it would seem that dentition-based harvest profiles can only be reliably computed for the combined caprine sample, even though they are likely to represent an amalgam of culling strategies that obscures any differences in exploitation of these two species.

Long-bone harvest profiles, in contrast, are usually based on larger samples of more reliably identified sheep and goat bones (Zeder and Lapham, in preparation), so that even smaller assemblages are likely to yield enough identifiable bones to generate species-level harvest profiles. In addition, the sequence and timing of long-bone fusion in sheep and goats has been shown to be more consistent and less influenced by differences in diet than are rates of tooth eruption and wear (Zeder, 2006a). Finally, long bones can be used to generate both species and sex-specific harvest profiles, providing a higher resolution picture of ancient exploitation strategies than species-level harvest profiles which represent an amalgam of male and female culling patterns (Zeder, 2001, 2005, 2008). The major limitation of long-bone harvest profiles is, of course, that they cannot capture the culling of older animals, whether these older animals are prime age wild males targeted by hunters, or older domestic females culled by herders once they pass peak reproductive years – ages that are captured in dentition-based harvest profiles.

We are not proposing that researchers reject one method for constructing caprine harvest profiles and embrace the other. We do argue, however, that the limitations and strengths of both long bone and dentition-based harvest profiles should be recognized. High resolution species-and sex-specific harvest profiles based on long bones are the most powerful tool we have for examining culling strategies directed at juvenile and early prime age animals (up to about five years of age). Taxonomically undifferentiated dental harvest profiles, on the other hand, seem best suited to examining generalized harvest strategies directed at older animals. Combining these two techniques to look at these two different phases of caprine harvest represents an optimal approach for gaining the most information about ancient exploitation strategies. The best way forward, then, is to acknowledge and control for the various limitations of both approaches while maximizing their different strengths in the computation and interpretation of harvest profiles of sheep and goats from archaeological sites.

Acknowledgements

The authors would like to thank the Department of Vertebrate Zoology, NMNH and the Department of Zoology, FMNH for access to collections used in this study; the NMNH Research Training Program for Pilaar's support while at the Smithsonian; and Marcia Bakry for her work on the figures. This paper has benefited from the thoughtful comments of Umberto Albarella, Jean-Denis Vigne, and one annonymous reviewer.

Appendix. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jas.2009.10.002.

References

Balaase, M., Ambrose, S.H., 2005. Distinguishing sheep and goats using dental morphology and stable carbon isotopes in C4 grassland environments. Journal of Archaeological Science 32, 691–702.

Boessneck, J., Müller, H.-H., Teichert, M., 1964. Osteologische Unterscheidungmerkmale zwischen Schaf (*Ovis aries* Linné) und Zeige (*Capra hircus* Linné). Kühn-Archiv 78, 1–29.

Boessneck, J., 1970. Osteological differences between sheep (Ovis ariea Linné) and Goats (*Capra hircus* Linné). In: Brothwell, D., Higgs, E. (Eds.), Science in Archaeology. Praeger, New York, pp. 331–358.

Bullock, D., Rackham, J., 1982. Epiphyseal fusion and tooth eruption of feral goats from Moffatdale, Dumfries and Galloway, Scotland. In: Wilson, B., Grigson, C., Payne, S. (Eds.), Ageing and Sexing Animal Bones from Archaeological Sites (BAR British Series 109). British Archaeological Reports, Oxford, pp. 55–72.

Deniz, E., Payne, S., 1982. Eruption and wear in the mandibular dentition as a guide to ageing Turkish Angora goats. In: Wilson, B., Grigson, C., Payne, S. (Eds.), Ageing and Sexing Animal Bones from Archaeological Sites (BAR British Series 109). British Archaeological Reports, Oxford, pp. 155–206.

Garcia-Gonzalez, R., 1983. Estudio de las ossificación postnatal en ovinos de raza rasa aragonesa. Munibe 33, 259–279.

Halstead, P., Collins, P., Isaakidou, V., 2002. Sorting the sheep from the goats: morphological distinctions between the mandibles and mandibular teeth of adult Ovis and Capra. Journal of Archaeological Science 29, 545–553.

Hatting, T., 1981. Osteological investigations on *Ovis aries* L. Dansk Naturhistorisk Forening 144, 115–135.

Helmer, D., 2000. Discrimination des genres Ovis et Capra à l'aide des prémolaires inférieures 3 et 4 et interpretation des ages d'abbatage: l'example de Dikili Tash (Grèce). Anthropozoologica 31/Ibex: Journal of Mountain Ecology 5, 29–38.

Hole, F., Flannery, K.V., Neely, J.A., 1969. Prehistory and Human Ecology on the Deh Luran Plain. Memoirs of the Museum of Anthropology, No. 1. The University of Michigan Press, Ann Arbor.

Jones, G.G., 2006. Tooth eruption and wear observed in live sheep from Buster Hill, the Cotswold farm park and five farms in the Pentland Hills, UK. In: Ruscillo, D. (Ed.), Recent Advances in Ageing and Sexing Animal Bones. Oxbow Books, Oxford, pp. 155–179.

Moran, N., O'Connor, T., 1994. Age attribution in domestic sheep by skeletal and dental maturation: a pilot study of available sources. International Journal of Osteoarchaeology 4, 267–285.

Noddle, B., 1974. Age of epiphyseal closure in feral and domestic goats and ages of dental eruption. Journal of Archaeological Science 1, 195–204.

Payne, S., 1973. Kill-off patterns in sheep and goats: the mandibles from Aşvan Kale. Anatolian Studies 23, 281–303.

Payne, S., 1985. Morphological distinctions between the mandibular teeth of young sheep, Ovis, and Goats, Capra. Journal of Archaeological Science 12, 139–147.

Payne, S., 1987. Reference codes for Wear Stages in the mandibular cheek teeth of sheep and goats. Journal of Archaeological Science 14, 609–614.

Prummel, W., Frisch, H.-J., 1986. A guide for the distinction of species, sex, and body size of sheep and goat. Journal of Archaeological Science 14, 609–614.

Redding, R., 1981. Decision Making in Subsistence Herding of Sheep and Goats in the Middle East. University Microfilms, Ann Arbor.

Silver, I., 1969. The ageing of domestic animals. In: Brothwell, D., Higgs, E. (Eds.), Science in Archaeology, second ed. Thames, London, pp. 283–302.

Vigne, J.-D., Helmer, D., 2007. Was milk a "secondary product" in the Old World Neolithisation process? Its role in the domestication of cattle, sheep and goats. Anthropozoologica 42, 9–40.

Zeder, M., 1991. Feeding Cities: Specialized Animal Economy in the Ancient Near East. Smithsonian Institution Press, Washington, D.C.

Zeder, M., 2001. A metrical analysis of a collection of modern goats (*Capra hircus aegagrus* and *Capra hircus hircus*) from Iran and Iraq: implications for the study of caprine domestication.

Zeder, M.A., 2005. New perspectives on livestock domestication in the fertile crescent as viewed from the Zagros Mountains. In: Vigne, J.-D., Peters, J., Helmer, D. (Eds.), The First Steps of Animal Domestication. Oxbow Books, Oxford, pp. 125–147.

Zeder, M.A., 2006a. Reconciling rates of long bone fusion and tooth eruption and wear in sheep (Ovis) and goat (Capra). In: Ruscillo, D. (Ed.), Ageing and Sexing Animals from Archaeological Site. Oxbow Press, Oxford, pp. 87–118.

Zeder, M.A., 2006b. A critical examination of markers of initial domestication in goats (*Capra hircus*). In: Zeder, M.A., Emshwiller, E., Smith, B.D., Bradley, D.G. (Eds.), Documenting Domestication: New Genetic and Archaeological Paradigm. University of California Press, Berkeley, pp. 181–208.

Zeder, M.A., 2008. Animal domestication in the Zagros: an update and directions for future research. In: Vila, E., Goucherin, L., Choyke, Buitenhuis, H. (Eds.), Archaeozoology of the Near East VIII. Travaux de la Maison de l'Orient et de la Méditerranée (TMO), Lyon, pp. 243–278.

Zeder, M.A., Lapham, H.A. Assessing the reliability of criteria used to identify postcranial bones of sheep, Ovis, and goats, Capra, in preparation.