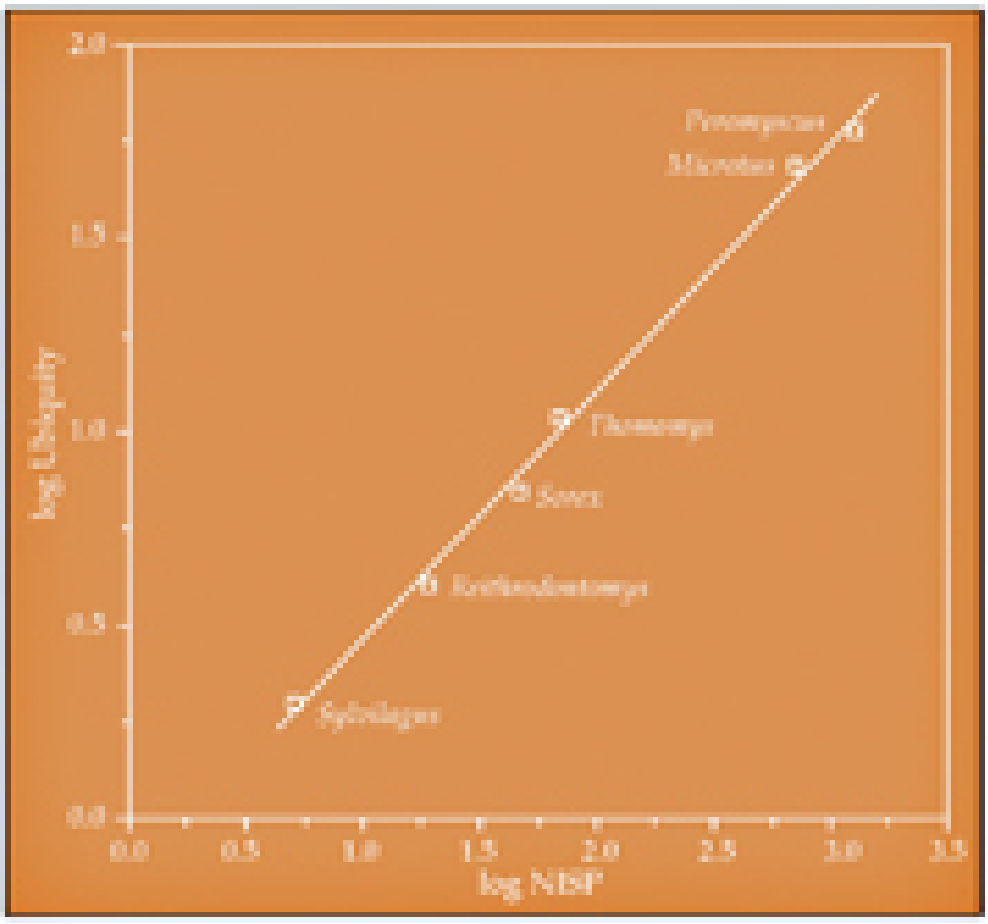
CAMBRIDGE MANUALS IN ARCHAEOLOGY

Quantitative Paleozoology

R. Lee Lyman



CAMBRIDGE www.cambridge.org/9780521887496

This page intentionally left blank

Quantitative Paleozoology

Quantitative Paleozoology describes and illustrates how the remains of long-dead animals recovered from archaeological and paleontological excavations can be studied and analyzed. The methods range from determining how many animals of each species are represented to determining whether one collection consists of more broken and more burned bones than another. All methods are described and illustrated with data from real collections, while numerous graphs illustrate various quantitative properties.

R. LEE LYMAN is professor of anthropology at the University of Missouri-Columbia. A scholar of late Quaternary paleomammology and human prehistory of the Pacific Northwest United States, he is the author of *Vertebrate Taphonomy*, and, most recently, the coeditor of *Zooarchaeology and Conservation Biology*.

Cambridge Manuals in Archaeology

General Editor

Graeme Barker, *University of Cambridge*

Advisory Editors

Elizabeth Slater, *University of Liverpool*

Peter Bogucki, *Princeton University*

Cambridge Manuals in Archaeology is a series of reference handbooks designed for an international audience of upper-level undergraduate and graduate students and for professional archaeologists and archaeological scientists in universities, museums, research laboratories, and field units. Each book includes a survey of current archaeological practice alongside essential reference material on contemporary techniques and methodology.

Books in the series

Pottery in Archaeology, CLIVE ORTON, PAUL TYERS, and ALAN VINCE
Vertebrate Taphonomy, R. LEE LYMAN
Photography in Archaeology and Conservation, 2nd edition, PETER G. DORRELL
Alluvial Geoarchaeology, A. G. BROWN
Shells, CHERYL CLAASEN
Sampling in Archaeology, CLIVE ORTON
Excavation, STEVE ROSKAMS
Teeth, 2nd edition, SIMON HILLSON
Lithics, 2nd edition, WILLIAM ANDREFSKY, JR.
Geographical Information Systems in Archaeology, JAMES CONOLLY and MARK LAKE
Demography in Archaeology, ANDREW CHAMBERLAIN
Analytical Chemistry in Archaeology, A. M. POLLARD, C. M. BATT, B. STERN, and S. M. M. YOUNG
Zooarchaeology, 2nd edition, ELIZABETH J. REITZ and ELIZABETH S. WING

Quantitative Paleozoology

R. Lee Lyman *University of Missouri-Columbia*

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK
Published in the United States of America by Cambridge University Press, New York
www.cambridge.org
Information on this title: www.cambridge.org/9780521887496

First published in print format 2008

ISBN-13 978-0-511-38646-6 eBook (EBL)

ISBN-13 978-0-521-88749-6 hardback

ISBN-13 978-0-521-71536-2 paperback

CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

Several years ago I had the opportunity to have a relaxed discussion with my doctoral advisor, Dr. Donald K. Grayson. In the course of that discussion, I asked him if he would ever revise his then 20-year-old book titled *Quantitative Zooarchaeology*, which had been out of print for at least a decade. He said "No" and explained that the topic had been resolved to his satisfaction such that he could do the kinds of analyses he wanted to do. A spur-of-the-moment thought prompted me to ask, "What if I write a revision?" by which I meant not literally a revised edition but instead a new book that covered some of the same ground but from a 20-years-later perspective. Don said that he thought that was a fine idea.

After the conversation with Grayson, I began to mentally outline what I would do in the book. I realized that it would be a good thing for me to write such a book because, although I thought I understood many of the arguments Grayson had made regarding the counting of animal remains when I was a graduate student, there were other arguments made by other investigators subsequent to the publication of Grayson's book that I didn't know (or if I knew of those arguments, I wasn't sure I understood them very well). I also knew that the only way for me to learn a topic well was to write about it because such a task forced me to learn its nuances, its underpinning assumptions, the interrelations of various aspects of the argument, and all those things that make an approach or analytical technique work the way that it does (or not work as it is thought to, as the case may be).

As I mentally outlined the book over the next several months, it occurred to me that at least one new quantitative unit similar to the traditional ones Grayson had considered had become a focus of analytical attention over the two decades subsequent to the publication of Grayson's book (MNE, and the related MAU). And an increasing number of paleozoologists were measuring taxonomic diversity – a term that had several different meanings for several different variables as well as being measured several different ways. What were those measurement techniques and

what were those measured variables? Finally, there were other kinds of phenomena that zooarchaeologists and paleontologists had begun to regularly tally and analyze. These phenomena – butchering marks, carnivore gnawing marks, rodent gnawing marks, burned bones – had become analytically important as paleozoologists had come to realize that to interpret the traditional quantitative measures of taxonomic abundances, potential biases in those measures caused by differential butchery, carnivore attrition, and the like across taxa had to be accounted for. As I indicate in this volume, there are several ways to tally up carnivore gnawing marks and the like, and few analysts have explored the fact that each provides a unique result.

Finally, it had become clear to me during the 1990s that many paleozoologists were unaware of what I took to be two critical things. First, zooarchaeologists seemed to seldom notice what is published in paleontological journals; at least they seldom referenced that literature. Thus, they were often ignorant of various suggestions made by paleontologists regarding quantitative methods. Paleontologists were equally unaware of what zooarchaeologists have determined regarding quantification of bones and shells and teeth. Therefore, it seemed best to title this volume *Quantitative Paleozoology* for the simple reason that were it to be titled "Quantitative Zooarchaeology," it likely would not be read by paleontologists. A very interesting book with the title *Quantitative Zoology* coauthored by a paleontologist (Simpson et al. 1960) already exists, so that title could not be used, aside from it being misleading. *Quantitative Paleozoology* is a good title for two reasons. The first reason is that the subject materials, whether collected by a paleontologist or an archaeologist, do not have a proximate zoological source (though their source is ultimately zoological) but rather have a proximate geological source, whether paleontological (without associated human artifacts) or archaeological (with associated and often causally related human artifacts). I conceive of all such remains as paleozoological. The second reason *Quantitative Paleozoology* is a good title is that the volume concerns how to count or tally, how to quantify zoological materials and their attributes, specifically those zoological remains recovered from geological contexts. Not all such topics are discussed here, but many are; for an introduction to many of those that are not, see Simpson et al. (1960), a still-useful book that was, fortunately, reprinted in 2003.

The second critical thing that many paleozoologists seem to be unaware of is basic statistical concepts and methods. I was stunned in 2004 to learn that an anonymous individual who had reviewed a manuscript I submitted for publication did not know what a "closed array" was and therefore did not understand why my use of this particular analytical tool could have been influencing (some might say biasing, but that is a particular kind of influencing) the statistical results. In the 1960s and early 1970s, many archaeologists and paleontologists did not have very high levels of statistical sophistication; I had thought that most of them did have such sophistication (or at

least knowledge of the basics) in the twenty-first century. The anonymous reviewer's comments indicate that at least some of them do not. Therefore, it seemed that any book on quantitative paleozoology had to include brief discussions of various statistical and mathematical concepts. In order to not dilute the central focus of the volume – quantitative analysis of paleozoological remains – I have kept discussion of statistical methods to a minimum, assuming that the serious reader will either already know what is necessary or will learn it as he or she reads the book. I have, however, devoted the first chapter to several critical mathematical concepts as well as some key paleozoological concepts.

Many of the faunal collections used to illustrate various points in the text were provided over the years by friends and colleagues who entrusted me with the analysis of those collections. Many of the things I have learned about quantitative paleozoology are a direct result of their trust. To these individuals, I offer my sincere thanks: Kenneth M. Ames, David R. Brauner, Jerry R. Galm, Stan Gough, Donald K. Grayson, David Kirkpatrick, Lynn Larson, Frank C. Leonhardy, Dennis Lewarch, Michael J. O'Brien, Richard Pettigrew, and Richard Ross. Perhaps more importantly, any clarity this book brings to the issues covered herein is a result of the collective demand for clarity by numerous students who sat through countless lectures about the counting units and methods discussed in this book. A major source of inspiration for the first several chapters was provided in 2004 by the Alaska Consortium of Zooarchaeologists (ACZ). That group invited me to give a daylong workshop on the topics of quantification and taphonomy, and that forced me to think through several things that had previously seemed less than important. I especially thank Diane Hansen and Becky Saleeby of the ACZ for making that workshop experience memorable.

An early draft of the manuscript was reviewed by Jack Broughton, Corey Hudson, Alex Miller, and an anonymous individual. Broughton and the anonymous reviewer ensured that a minimum of both glaring errors in logic and stupid errors in mathematics remain in this version. Broughton and the anonymous reviewer insisted that I include several recently described analytical techniques, and they identified where I overstepped and where I misstepped. These individuals deserve credit for many of the good things here.

I wrote much of the first draft of this volume between July 2005 and August 2006. During that time, I lost my younger brother and both parents. They all had an indirect hand in this book. My parents taught me to hunt and fish, and all of the things that accompany those activities. My brother did not discourage me from collecting owl pellets from his farm equipment shed, or laugh too hard when I collected them; he even grew to appreciate what could be learned from the mouse bones they contained. I miss them all, and I dedicate this book to the three of them.

June 2007

1

Tallying and Counting: Fundamentals

Early in the twentieth century, paleontologist Chester Stock (1929) was, as he put it, faced with "recording a census" of large mammals from the late Pleistocene as evidenced by their remains recovered "from the asphalt deposits of Rancho La Brea," in Los Angeles, California. Paleontologist Hildegarde Howard (1930) was faced with a similar challenge with respect to the bird remains from Rancho La Brea. Stock and Howard could have merely listed the species of mammals and the species of birds, respectively, that were represented by the faunal remains they had – they could have constructed an *inventory* of taxa – but they chose to do something more informative and more analytically powerful. They tallied up how many individuals of each species were represented by the remains – they each produced a census. The quantitative unit they chose became known as the *minimum number of individuals*, or MNI, a unit that was quickly (within 25 years) adopted by many paleozoologists. We will consider this unit in some detail in Chapter 2, but here it is more important to outline how Stock and Howard defined it and why they decided to provide a census rather than an inventory of mammals and an inventory of birds.

Stock (1929:282) stated that the tally or "count" of each taxon was "determined by the number of similar parts of the internal skeleton as for example the skull, right ramus of mandible, left tibia, right scaphoid. In many cases the total number of individuals for any single group [read *taxon*] is probably a minimum estimate." Howard (1930:81–82) indicated that "for each species, the left or the right of the [skeletal] element occurring in greatest abundance was used to make the count. . . . It is probable that in many instances the totals present a minimum estimate of the number of individuals [per taxon] actually represented in the collection." We will explore why the procedure Stock and Howard used provides a "minimum" estimate of abundance in Chapter 2. Stock and Howard each produced a type of pie diagram to illustrate their respective censuses of mammalian and of avian creatures based on the bony remains of each (Figure 1.1).

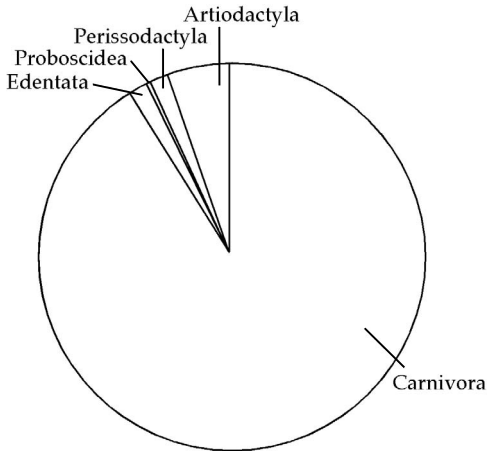


FIGURE 1.1. Chester Stock's pie diagram of abundances of five mammalian orders represented in faunal remains from Rancho La Brea. Redrawn from Stock (1929).

An inventory of the mammalian taxonomic orders Stock identified among the bones and teeth he studied would look like this:

Carnivora
Edentata
Proboscidea
Perissodactyla
Artiodactyla

Clearly, the pie diagram in Figure 1.1 reveals more about the structure of the Rancho La Brea mammalian fauna because it contains not only the same set of taxonomic orders as the inventory, but it also contains measures of the abundances of animals belonging to each order. This example illustrates one of the major reasons why paleozoologists count or tally the animal remains they study. Taxa present in a collection can, on the one hand, be treated as attributes or as present or absent from a fauna, such as is given in the inventory above (sometimes referred to as a "species list" if that taxonomic level is used). On the other hand, abundances of each taxon provide a great deal more information about the prehistoric fauna. There are times when knowing only which taxa are present, or knowing only what the frequencies of different taxa are is all that is wanted or needed analytically. (Two faunas may have the same, or quite different, frequency distributions of individual organisms across taxa, and the research question may only require knowing the frequency distributions and not the taxa.) Knowing both, however, means we know more than when we know just one or the other. And that is a good reason to count faunal remains and to determine

a census. Counting faunal remains, particularly old or prehistoric remains, and the variety of attributes they display, whether the remains are from archaeological or paleontological contexts, is what this book is about.

There is already a book about counting animal remains recovered from archaeological and paleontological sites (Grayson 1984), and several other volumes cover some of the same ground, if in less detail (e.g., Hesse and Wapnish 1985; Klein and Cruz-Uribe 1984; Reitz and Wing 1999). Noting this, one could legitimately ask why another book on this topic is necessary. There are several reasons to write a new book. Much has happened in the field since Grayson (1984) published his book (and his book has been out-of-print for several years). Some of what has happened has been conceptually innovative, such as the definition of new quantitative units meant to measure newly conceived properties of the paleozoological record. Some of what has happened has been technically innovative, such as designing new protocols for tallying animal remains that are thought to provide more accurate reflections of what is represented by a collection of remains than tallies based on less technologically sophisticated methods. And, some of what has happened is misguided or archaic, such as arguing that if certain biological variables are not mathematically controlled for, then any count of taxonomic abundances is invalid. It is time (for these reasons) for a new, up-to-date examination of the quantitative units and counting protocols paleozoologists use in their studies.

There is yet another reason to produce a new book on quantitative paleozoology. Today, early in the third millennium, there are more people studying paleozoological collections than there were 20 years ago. These folks need to be able to communicate clearly and concisely with one another regarding their data and their analyses because the use of ambiguous terminology thwarts efficient communication and results in confusion. This point was made more than a decade ago with respect to the plethora of terms, many unfamiliar to those in the field, used for quantitative units in zooarchaeology (Lyman 1994a). Yet, the problem continues today. This problem had originally been identified more than 15 years earlier still by Casteel and Grayson (1977). For whatever reason, terminological ambiguity seems to plague paleozoology and continues to do so despite it being explicitly identified twice in the past 30 years.

In my earlier discussion of terminological ambiguity (Lyman 1994a), I did not advocate a particular terminology, nor am I doing so here. Clearly there are terms I prefer – the ones I use in this volume are the ones I learned as a student. What I am arguing here is that whatever terms or acronyms one uses, these must be clearly defined at the start so as to avoid misunderstanding. In reading and rereading the literature on quantitative paleozoology as I prepared this book, I was often dumbfounded when people used terms such as "bone" and "relative abundance" when it

was quite clear that they were discussing teeth and absolute abundances, respectively. Much of the remainder of this chapter is, therefore, devoted to terminology and definitions. For quick reference, I have included a glossary of key terms at the end of this volume.

In this introductory chapter, several basic mathematical and statistical concepts are defined. This is necessary because these concepts will be used throughout subsequent chapters and thus the concepts must be understood in order to follow the discussion in later chapters. Several basic paleozoological concepts are introduced and defined for the same reason. I begin with these concepts before turning to the mathematical and statistical concepts.

PALEOZOOLOGICAL CONCEPTS

Throughout this volume the focus is on vertebrates, especially mammals, because that is the taxonomic group which much of the literature concerns and because it is the group with which I am most familiar. However, virtually every thing that is said about quantifying vertebrate remains and their attributes holds with equal force for invertebrates (e.g., Claassen 1998:106–107).

In many discussions of how paleofaunal remains are tallied, and even in some discussions of how modern animal bones should be counted, the reader may encounter the term "skeletal element." Or, one might encounter the term "bone," or "tooth," or "shell," or any of many other similar, more or less synonymous general terms for skeletal remains. But if one collection comprises ten "bones" of a skeleton and another consists of eleven "bones" of another skeleton of the same species as the first, is the latter more anatomically complete than the former? Is the taxon less abundant in the first collection than in the second? If you think the answer is "Yes" to either question, you might be correct. But you could be wrong if when the analyst tallied specimens no distinction was made between anatomically complete bones and fragments of bones. The lesson is simple. If we are going to tally up skeletal parts and want to compare our tally with that of another analyst working with another collection, we had best be sure that we counted skeletal parts the same way that the other person did. What, then, exactly is a skeletal element?

Paleontologist Michael Voorhies (1969:18) distinguished between "fragments" and "elements or bones," but we need something more explicit and inclusive because not all skeletal elements are, technically, bones. Some are teeth, some are horns, and some are antlers, and so on. Following Arnold Shotwell (1955, 1958), Donald Grayson (1984) and Catherine Badgley (1986) provide useful terminology and definitions. A

skeletal element is a complete discrete anatomical unit such as a bone, tooth, or shell. The critical phrase is *complete discrete anatomical unit*. Each such item is a discrete "anatomical organ" (Francillon-Vieillot et al. 1990:480) that does not lose its integrity or completeness when it is removed from an organism. A humerus, a tibia, a carpal, a first lower molar – each is a skeletal element. One might correctly note that "discreteness" depends on the age or ontogenetic stage of development of the organism, but many paleozoologists would not tally the proximal epiphysis of a humerus and the diaphysis of that humerus as two separate specimens if it was clear that the two specimens went together (an issue we return to in Chapter 2). Those same paleozoologists usually don't tally up each individual tooth firmly set in a mandible, along with the dentary or mandible bone. These are potentially significant concerns but may ultimately be of minimal analytical import once we get into tallying specimens.

Not all faunal remains recovered from paleozoological deposits are anatomically complete; some are represented by only a part of the original skeletal element because of fragmentation. Thus, another term is necessary. A *specimen* is a bone, tooth, or shell, or fragment thereof. All skeletal elements are specimens, but not all specimens are skeletal elements. A distal humerus, a proximal tibia, and a fragment of a premolar are all specimens that derive from skeletal elements; phenomenologically they are not, technically, anatomically complete skeletal elements. *Specimen* is an excellent term for many counting operations because it is value-free in the sense that it does not reveal whether specimen A is anatomically more complete, or less complete, than specimen B. We can record whether specimen A is anatomically complete, and if it isn't, we can record the portion of a complete element that is represented by a fragment, if our research questions demand such. *Specimen* is also a better generic term than *skeletal element* for the individual skeletal remains we study because *skeletal element* implies that a complete anatomical unit is represented. The problem with the terms "bone" and "tooth" and the like are that sometimes when analysts use them they mean both anatomically complete skeletal elements as defined here and incomplete skeletal elements. Failure to distinguish the two kinds of units – skeletal element and specimen – can render separate tallies incomparable and make the significance of various analyses obscure. Throughout this volume, I use the term *skeletal part* as a synonym for *specimen*, but whereas the latter is a general category that can include many and varied anatomical portions, skeletal part is restricted to a particular category of anatomical portion, say, distal humerus. *Skeletal portion* is sometimes used in the same category-specific way that skeletal part is but will usually mean a multiple skeletal element segment of a skeleton, such as a forelimb.

Henceforth, in this volume, *specimen* will be used to signify any individual skeletal remain, whether anatomically complete or not. Unfortunately, the terms "skeletal

Table 1.1. *An example of the Linnaean taxonomy*

Taxonomic level	Taxonomic name	Common name
Kingdom	Animalia	Animals
Phylum	Vertebrata	Vertebrate
Class	Mammalia	Mammals
Infraclass	Eutheria	Placental mammal
Order	Carnivora	Carnivores
Family	Canidae	Canids
Genus	*Canis*	Dogs, coyotes, wolves, and allies
Species*	*latrans*	coyote

*Technically, the species name is *Canis latrans*; *latrans* is the specific epithet.

element" and "element" are still often used to denote anatomically incomplete items. An effort is made throughout this book to make clear what exactly is being tallied and how it is being tallied. In this respect, what are usually tallied are what are termed "identified" or "identifiable" specimens. Typically, this means identified as to biological taxon, usually genus or species, represented by a bone, tooth, or shell (Driver 1992; Lyman 2005a). To identify skeletal remains, one must know the structure of the Linnaean taxonomy, an example of which is given in Table 1.1 . One must also know the basics of skeletal anatomy, by which is meant that one must know the difference between a scapula and a radius, a femur and a cervical vertebra, a clavicle and a rib, and so on. Finally, the person doing the identifications must be able to distinguish intertaxonomic variation from intrataxonomic variation. *Intrataxonomic variation* is also sometimes termed "individual variation" within the species level of the taxonomy. I presume that readers of the book know these things, along with anatomical location and direction terms used in later chapters.

The importance of the requirements for identification should be apparent when one realizes that "identification" involves questions such as: Is one dealing with a mammal or a bird? If it is a mammal, is it a rodent or a carnivore? If it is a carnivore, is it a canid, a felid, a mustelid, or any of several other taxa of carnivores? The importance of the other knowledge requirement – basic skeletal anatomy – will assist in answering the questions just posed. The importance of distinguishing intertaxonomic from intrataxonomic variation is usually (and best) met by consultation of a comparative collection of skeletons of known taxonomic identity. The procedure is simple. Compare the taxonomically unknown paleozoological specimen with comparative specimens of known taxonomy until the best match is found. Often the closest match will be obvious, and the unknown specimen is "identified" as belonging to the same taxon as the known comparative specimen. Sometimes this means that

one may be able to determine the species represented by the paleozoological specimen, but other times only the genus or perhaps only the taxonomic family or order will be distinguishable.

Taxonomic identification is a complex matter that is discussed at length in other contexts (e.g., Driver 1992; Lyman 2005a; and references therein). Blind tests of identification results (e.g., Gobalet 2001) highlight the practical and technical difficulties. For one thing, what is "identifiable" to one analyst may not be to another (e.g., Grayson 1979). Gobalet (2001) provides empirical evidence for such interanalyst variation. It is precisely because of such interobserver differences and the interpretive significance of whether, say, a bone is from a bobcat (*Lynx rufus*) or a North American lynx (*Lynx canadensis*) that paleontologists developed a standardized format for reporting their results. Specimens (not necessarily anatomically incomplete skeletal elements) are illustrated and are verbally described with taxonomically distinctive criteria highlighted so that other paleontologists can independently evaluate the anatomical criteria used to make the taxonomic identification. Zooarchaeologists have been slow to understand the importance of this reporting form (see Driver [1992] for a noteworthy exception). This is not the place to delve further into the nuances of taxonomic identification and how to report and describe identified specimens. What is important here is to note that skeletal remains – faunal specimens – are usually tallied by taxon. "There are X remains of bobcats and Y remains of lynx." So, identification must precede tallying. To make taxonomic identifications, one must first determine which skeletal element is represented by a specimen is order to know whether the paleozoological unknown should be compared to femora, humeri, tibiae, and so on. And sometimes the frequencies of each skeletal element or each part thereof are analytically important.

The final paleozoological concept that requires definition is *taphonomy*. The term was originally coined by Russian paleontologist I. A. Efremov (1940:85) who defined it as "the study of the transition (in all details) of animal remains from the biosphere into the lithosphere." Although not without precedent, Efremov's term is the one paleozoologists (and an increasing number of paleobotanists) use to refer to the processes that influence the creation and preservation (or lack thereof) of the paleobiological record. We will have reason to return again and again to this basic concept; here it suffices to note that a taphonomic history concerns the formation of an assemblage of faunal remains. Such a history begins with the accumulation and deposition of the first specimen, continues through the deposition of the last specimen, through the preservation, alteration, and destruction of remains, and up to collection of a sample of the remains by the paleozoologist (see Lyman [1994c] for more complete discussion). Along the way, faunal remains are modified, broken, and even destroyed. The modification, fracture, and destruction processes create

and destroy different kinds of phenomena the observation of which can generate quantitative data.

A final note about how paleozoological data are presented in the book. Capital letters are used to denote upper teeth, lower case letters to denote lower teeth, and a lowercase d to denote deciduous premolars. Thus, a permanent upper second premolar is P2, a deciduous lower third premolar is dp3, and a lower first molar is m1. The capital letter L is used to signify the left element of bilaterally paired bones, and the capital letter R is used to signify the right element. In general, D stands for distal, and P stands for proximal. The critical thing to remember is the difference between a *specimen* and a *skeletal element*; both terms will reappear often in what follows, and both kinds of units can be identified and tallied.

MATHEMATICAL AND STATISTICAL CONCEPTS

This book is about quantification, but the topics covered include different sorts of quantification, particularly counting or tallying units, methods of counting, and analyzing counts. A term that might have been used in the title of the book, were it not for its generality, is *measurement*. Typically this term is defined as assigning a numerical value to an observation based on a rule governing the assignment. The rule might be that length is measured in linear units of uniform size, such that we can say something like "Pencil A is 5 cm long and pencil B, at 10 cm of length, is twice as long as pencil A." *Measurement* more generally defined concerns writing descriptions of phenomena according to rules. An *estimate* is a measurement assigned to a phenomenon (making a measurement) based on incomplete data. The process of *estimation* can involve judging how tall someone is in centimeters without the benefit of a tape measure, or studying a flock of birds and suggesting how many individuals there are without systematically tallying each one. Making estimates, like taking measurements, is a way to describe phenomena. Descriptions involve attributes of phenomena that may or may not have numerical symbols or values associated with them. Whether they do or not concerns what is often referred to as the scale of measurement of the attribute that is under scrutiny.

Scales of Measurement

Stock's census of Rancho La Brea mammals (Figure 1.1) illustrates that quantitative data describing taxonomic abundances are more revealing than taxonomic

presence–absence data. Quantitative data often are subjected to a variety of mathematical manipulations and statistical analyses. Those manipulations and analyses are only valid if the data are of a certain kind. Four distinct scales of measurement are often distinguished (Blalock 1960; Shennan 1988; Stevens 1946; Zar 1996), and it is important that these be explicitly defined at the start because they will be referred to throughout the book.

Nominal scales of measurement are those that measure differences in kind. Of the several scales they contain the least amount of information. Numbers may be assigned to label nominal scale phenomena, such as 0 = male, 1 = female; or 11 = quarterback, 32 = fullback, and 88 = wide receiver on a football team. Or, numbers need not be assigned, but rather labels used such as Italian citizen, French citizen, and German citizen; or coyote (*Canis latrans*), wolf (*Canis lupus*), and domestic dog (*Canis familiaris*). Nominal scales of measurement do not include an indication of magnitude, ordering, or distance between categories, and are sometimes labeled *qualitative attributes* or *discontinuous variables*. They are qualitative because they record a phenomenon in terms of a quality, not a magnitude or an amount. They are discontinuous (or discrete) because it is possible to find two values between which no other intermediate value exists; there is (normally) no organism that is halfway between a male and a female within a bisexual species. Other scales of measurement tend to be quantitative because they specify variation more continuously. *Continuous variables* are those that can take any value in a series, and there is always yet another value intermediate between any two values. A tally of skeletal specimens of coyote in an archaeological collection may be 127 or 128, but there won't be a collection in which there are 127.5 or 127.3 or 127.924 specimens of coyote. But the lengths of coyote humeri are continuous; think about the numbers just noted as millimeters of length.

Ordinal scales of measurement are those that record greater than, less than relationships, but not the magnitude of difference in phenomena. They allow phenomena to be arranged in an order, say, from lesser to greater. "I am older than my children" is a statement of ordinal scale difference, as is "The stratum on the bottom of the stratigraphic column was deposited before the stratum on the top of the column" and "A year is longer than a month." There is no indication of the magnitude of difference in my age and the ages of my children, or in the length of time between the deposition of the bottom and top strata, nor in the duration of a year relative to the duration of a month. Instead, we only specify which phenomenon is older (or younger), or which was deposited first (or last), or which is longer in duration (or shorter). Sometimes when one uses an ordinal scale, measurements are said to be *relative* measurements because a measure of phenomenon A is made relative to

phenomenon B; A is older/shorter/heavier than B. Ordinal scale measurements may be (and often are said to be) *rank ordered* from greatest to least, or least to greatest, but the magnitude of distance between any two measurements in the ordering is unknown. Ordinal scale measurements are discrete insofar as there is no rank of "first and a half" between the rank of first and second (ignoring tied ranks).

Interval scales of measurement are those that record greater than or less than relationships and the magnitude of difference between phenomena. Both the order of measurements and the distance between them are known. My children are 23 and 25 years old; I am 56 years old, so I am 33 and 31 years older than my two children, respectively. The stratum on the bottom of the stratigraphic column has an associated radiocarbon date of 3000 BP and the stratum on top has an associated date of 500 BP, so the stratum on the bottom was deposited about 2,500 (^{14}C) years before the stratum on top (assuming the dated materials in each stratum were formed and deposited at the same time as the strata were deposited). On average, a year is 365.25 days long whereas an average month is about 30.4 days in duration; the difference in duration of an average year and an average month is thus 334.85 days. The distance between 10 and 20 units (days, years, centimeters) is the same as the distance between 244 and 254 of those units, the same as the distance between 5337 and 5347 of those units, and so on. Interval scales are typically used to measure what are referred to as *quantitative variables*. Interval scale measurements, like ordinal scale ones, can be rank ordered from greatest to least, or least to greatest, but unlike with ordinal scale measures, the distance between any two interval scale measurements is known. Indeed it must be known else the variable is not interval scale. Interval scale measurements are generally continuous but may be discrete. If age is recorded only in whole years, then age is continuous but it is also discrete (ignoring for the sake of discussion that one might be 53.7 years old). Importantly, interval scale measures have an arbitrary zero point. It can be 0°Celsius outside, but there is still heat (if seemingly only a little) caused by the movement of molecules. The zero point on the Celsius scale is placed at a different location along the continuum of amount of molecular movement than is the zero point of the Fahrenheit temperature scale. Both zero points are arbitrary with respect to the amount of heat (molecular movement), thus both measures of temperature are interval scale.

Ratio scales of measurement are identical to interval scales but have a natural zero. Thus, the theoretical natural zero of temperature is –273° Celsius (or 0 Kelvin, or –459° Fahrenheit). There is no molecular movement at that temperature. Similarly, a mammal in a cage comprises 1 individual consisting of more than 100 bones and teeth, but if the cage is empty there are 0 (zero) individuals, 0 bones, and 0 teeth in the cage. Thus, if a taxon is represented by 0 skeletal specimens in an assemblage, it is

absent; that is a natural zero. Essentially all quantitative measures in paleozoology – taxonomic abundances, frequency of gnawing damage, and so on – are potentially ratio scale. Whether they are in fact ratio scale or not is another matter.

Measurements of different scales allow (or demand, depending on your perspective) different statistical tests of different scales or power. Thus, ordinal scale measurements require ordinal scale statistical tests; interval/ratio scale measurements can be analyzed with either interval/ratio scale statistics or ordinal scale ones, but the reverse – applying interval/ratio scale statistics (or parametric statistics) to ordinal scale data – will likely violate various statistical assumptions.

Most paleozoologists working in the twentieth century sought ratio scale measures of the attributes of the ancient faunal remains they studied, just as Stock and Howard did. Although the optimism that such measures would eventually be designed has waned somewhat, there are still many who hope for such, whether working with human remains (e.g., Adams and Konigsberg 2004), paleontological materials (e.g., Vermeij and Herbert 2004), or zooarchaeological collections (e.g., Marean et al. 2001; Rogers 2000a). We now know a lot more about taphonomy than we did even 20 years ago when Grayson (1984), Klein and Cruz-Uribe (1984), and Hesse and Wapnish (1985) noted that many problems with quantitative zooarchaeology originated in taphonomic histories. And we also know that many taphonomic analyses and interpretations of taphonomic histories require quantitative data and analyses of various sorts. Where taphonomy can influence quantitative paleozoology is noted throughout this volume, and it is occasionally suggested what we might do about those influences. The point here is that ratio scale measurements of faunal remains and many of their attributes may be precluded because of taphonomic history.

Measured and Target Variables: Reliability and Validity

Other important statistical concepts concern the difference between a *measured variable* and a *target variable*. A measured variable is what we actually measure, say, how many gray hairs I have on my head. A target variable is the variable that we are interested in, say, my age. The critical question is this: Are the measured variable and the target variable the same variable, or are they different? If the latter, the question becomes: Are the two variables sufficiently strongly correlated that measuring one reveals something about the other? It is likely that the number of gray hairs on my head will be correlated with my age, assuming I do not artificially color my hair (either not gray, or gray). But although the color of the shirt I am wearing today

can be measured rather precisely, it is unlikely to indicate or correlate with my age (although the style of my shirt might).

The concepts of measured variable and target variable can be stated another way. When we measure something, are we measuring what we think we are measuring? Does the attribute we are measuring reflect the concept (e.g., length, age, color) we wish to describe (Carmines and Zeller 1979)? These questions serve to define the concept of *validity*. Is a radiocarbon age on a piece of burned wood a *valid* measure of the age of deposition of a fossil bone with which the wood is stratigraphically associated? Assuming no contamination of the sample of wood, and that the wood was deposited more or less simultaneously with the bone, it will be a valid measure if it derives from a plant that was alive at about the same time as the animal represented by the bone. Validity is a different property of a measurement than *reliability*, which simply defined means replicability, or, if we measure something twice, do we get the same answer? If, on the one hand, we measure the length of a femur today and get 12.5 cm, tomorrow we measure it and get 12.4 cm, and the next day we measure it and get 12.5 cm, then we are producing rather consistent and thus reliable measures of that femur's length. On the other hand, femur length is unlikely to be a valid measure of the time period when the represented animal was alive, regardless of the reliability of our measurements of length.

Another set of measurements will help underscore the significance of the preceding paragraph, and help highlight the differences between a *target variable* and a *measured variable*. A *fundamental measurement* (sometimes referred to as *primary data* [Clason 1972; Reitz and Wing 1999]) is one that describes an easily observed property of a phenomenon. Length of a bone, stage of tooth eruption in a mandible, and taxon represented by a shell are all fundamental measurements. A *derived measurement* (sometimes referred to as *secondary data*) is more complex than a fundamental one because it is based on multiple fundamental measurements. Derived measurements are defined by a specified mathematical (or other) relation between two or more fundamental measurements. A ratio of length to width exemplifies a derived measure. Derived measurements require analytical decisions above and beyond a choice of scale; do we calculate the ratio of length to width, or width to length, or width to thickness? As a result, derived measurements are sometimes difficult to relate clearly to theoretical or interpretive concepts. Derived measurements may nevertheless reveal otherwise obscure patterns in data even though relating those patterns to a target variable may be difficult.

The MNI measure mentioned above is the most widely known derived measurement in paleozoology. It depends on (i) tallies of (ii) each kind of skeletal element of (iii) each taxon in a collection, and often (iv) (but not always) other information, such as size of bones of a taxon. Each of the lower case Roman numerals denotes

a distinct fundamental measurement; each plays a role in deriving an MNI, as can several other fundamental measurements (considered in more detail in Chapter 2). A *fiat or proxy measurement* will likely be more complex than either a fundamental or derived measurement because a fiat measurement is more conceptual or abstract and less easily observed. The distinction of fundamental, derived, and proxy measurements is relevant to a measurement's accuracy. "Accuracy" refers to "the nearness of a measurement to the actual value of the variable being measured" (Zar 1996:5). Throughout this volume, major concerns are the accuracy and validity of derived measures or secondary data, and fundamental measures or primary data with respect to a target variable of interest. Does a particular derived measure, such as MNI, accurately reflect the abundance of individual organisms in a collection of bones and teeth (or shells)? Of organisms in a deposit? Of organisms on the landscape?

Stock's census of Rancho La Brea mammals was, he hoped, an accurate proxy measure of the structure and composition of the mammalian fauna on the landscape at the time of the deposition of the remains. That long-dead fauna is not directly visible or measurable, so how well the remains from the tar pits actually reflect or measure that fauna in terms of which taxon was most abundant and which was least abundant and a host of other properties (how accurately MNI measures the landscape fauna) cannot be determined. The validity of a fiat or proxy measurement, or a measured variable, for reflecting a target variable of some sort is the key issue underpinning much of the discussion in this volume. This is so for the simple reason that many target variables in paleozoology cannot be directly measured reliably or validly with broken bones, isolated teeth, and fragments of mollusk shell. What this book is in part about is how well the measured variables and proxy measurements commonly used by paleozoologists measure or estimate the target variable(s) of interest. Two key questions to keep in mind throughout this book are: What is the target variable? How is the measured variable related to the target variable of interest? As a prelude to how important these questions are, think about this. Was Stock wise to use MNI (the derived and measured variable) to estimate the abundances of mammals *on the landscape* (the target variable) given that he only had animals that became mired in the pits of sticky tar at Rancho La Brea? Would he have been better off using, say, the tally of skulls (a different measured variable) to estimate the abundances of mammals *trapped in the tar pits* (a different target variable)?

Absolute and Relative Frequencies and Closed Arrays

An *absolute frequency* is a raw tally of some set of entities, usually all of a particular kind. To note that there are ten rabbit bones and five turkey bones in a collection is

to note the absolute frequencies of specimens of each species. If one were to note that in that collection of fifteen specimens, 66.7 percent of the specimens were of rabbits and 33.3 percent were of turkey, then one would be noting the *relative frequency* of each species. Relative frequencies are termed such because they are *relative* to one another. A relative frequency is a quantity or estimate that is stated in terms of another quantity or estimate. The analyst could have different absolute abundances, say thirty rabbit bones and fifteen turkey bones, but rabbit bones would comprise the relative abundance of 66.7 percent of the collection and turkey bones would comprise 33.3 percent of that collection, the same as when there are ten rabbit bones and five turkey bones. Percentages and proportions of a total are relative frequencies. The term "relative frequencies" is sometimes used in the paleozoological literature to signify estimates in which a quantity is not stated but rather only that A is greater (or smaller, or less) than B. In such cases relative frequencies are equivalent to ordinal scales of measurement. In this volume, the term "relative frequencies" is used in the more typical sense of percentage or proportional abundances.

Relative frequencies are typically given as percentages of some total set of things, and the summed relative frequency is always 100 percent (proportions are fractions). When relative frequencies of kinds of things in a set of things are given as percentages, all of those frequencies must sum to 100 percent rather than 90 percent or 110 percent. Such percentage relative frequencies comprise what is called a *closed array* (proportions also form a closed array as they must sum to 1.0).

Another way to think about the difference between absolute and relative frequencies involves comparison of measurements. Let's say we have two collections of faunal remains. In collection 1, taxon A is represented by 5 specimens and taxon B is represented by 10 specimens. In collection 2, taxon A is represented by 50 specimens and taxon B is represented by 55 specimens. The absolute difference in abundances of the two taxa in each collection is 5 specimens, but in collection 1, taxon A is only 50 percent as abundant as taxon B whereas in collection 2 taxon A is 90.9 percent as abundant as taxon B. Or, one could say that in collection 1 the relative abundances of taxa A and B are 33.3 percent and 66.7 percent, respectively, whereas the relative abundances of those taxa in collection 2 are 47.6 percent and 52.4 percent, respectively. The difference between absolute and relative frequencies is not a matter of which is correct and which is not, but rather they are simply two different ways to measure (describe) the frequencies of things.

Importantly, the absolute frequency of things of kind A in a collection will not change value if the absolute frequency of kind B in that collection changes, but the relative frequency of both A and B will change if the absolute frequency of either A or B changes. This last property is a characteristic – one could say diagnostic – of

Table 1.2. *Fictional data on the absolute abundances of two taxa in six chronologically sequent strata*

	Taxon A	Taxon B
Stratum VI	50 (71.4)	20 (28.6)
Stratum V	50 (62.5)	30 (37.5)
Stratum IV	50 (55.6)	40 (44.4)
Stratum III	50 (50.0)	50 (50.0)
Stratum II	50 (45.4)	60 (54.6)
Stratum I	50 (41.7)	70 (58.3)

Relative (percentage) abundances in parentheses.

closed arrays; they must sum to 100 percent. Consider the set of fictional data in Table 1.2. If we examine these data, we see that Taxon A does not change in absolute abundance over the stratigraphic sequence, but Taxon B does change in absolute abundance. However, relative abundance data suggest that both taxa change in abundance. This example reveals a final and critically important aspect of abundance data.

In paleozoology, absolute abundance data or raw tallies are often given, but when it comes to interpreting abundance data, it is in terms of relative abundances. Using the fictional data in Table 1.2 as an example, one might read something like the following:

> Throughout the stratigraphic sequence (from stratum I as earliest or oldest to stratum VI as youngest) Taxon A increased in abundance relative to Taxon B, which decreased in relative abundance. Given that Taxon A prefers habitats that support vegetation adapted to cool-moist climatic conditions, and Taxon B prefers habitats indicative of warm-dry climatic conditions, then it seems that over the time span represented by strata I–VI, the local climate became progressively cooler and moister.

Notice that in the interpretation no mention is made of the absolute abundances of taxa A and B. Rather, their abundances *relative to each other* are the focus. The abundances are not even taken as measures of how many of either taxon was present on the landscape at the time the strata and faunal remains were deposited. Rather, the interpretation involves postulating a cause for the shift in relative abundances of two taxa. One could also postulate that hunting practices or procurement technology shifted, resulting in the shift in which taxon was taken more frequently. Resolving these sorts of issues is beyond the scope of this volume, but suffice it to say that regardless of the interpretive model one calls upon, *relative* abundances, in the case of this example relative *taxonomic* abundances, are interpreted.

DISCUSSION

Quantitative data often comprise tallies of different kinds of phenomena. They might also include a set of measurements of, say, the length of individual specimens. This book concerns only the former kind of quantitative data – tallies. It is about how a paleozoologist might count phenomena (faunal specimens, or attributes thereof) when one seeks a measure of the magnitude of a particular variable that demands counts of phenomena (bones, teeth, shells, and fragments thereof, or burned bones, gnawed bones, or broken bones). How one chooses to tally those phenomena, and how those tallies are summarized and analyzed statistically, depend in large part on the research question asked and the target variable that one hopes to measure in order to answer that question. The choices likely will also depend on the presumed relationship of the target variable and the chosen measured variable. Discussion of how one determines the nature of that relationship in particular cases is beyond the scope of this volume. When necessary to assist discussions in later chapters, a particular relationship is assumed or identified.

The terms *assemblage* or *collection* denote an aggregate of faunal remains whose setness has been defined archaeologically (e.g., remains from an excavation unit), geologically (e.g., remains from a trash pit or a stratum), or analytically (e.g., all remains of a taxon). Graphs are used whenever possible to exemplify and illustrate concepts and analytical results, and to display relationships between variables. Statistical analyses are kept relatively simple and are used to evaluate particular properties of collections. In a few cases, statistical complexities are described in a clearly delineated box of text and may be skipped when reading the main content of a chapter. Data are often presented in table form so that the reader may replicate the statistical analyses (and graphs) to ensure understanding. This volume is, however, not meant to be exhaustive with respect to all of the myriad ways that faunal remains might be counted, or with respect to how the varied features faunal remains might display can be counted. Rather, most of the commonly used quantitative units (measured variables) and their attendant analyses serve as the background against which the discussion is framed. Target variables are identified and defined as necessary when discussing particular quantitative units.

This is not a book about taphonomic, zooarchaeological, or paleontological analyses. There are several excellent titles on each of these topics that are presently available (e.g., Lyman 1994c; Reitz and Wing 1999; Simpson et al. 1960, respectively). *Quantitative Paleozoology* is meant as a supplement to those other volumes because it covers in detail a limited range of topics relevant to various analytical methods and techniques described in each of those other volumes.

BACKGROUND OF SOME FAUNAL SAMPLES

Throughout this volume, extensive use is made of data derived from actual zooarchaeological and paleontological collections of vertebrate, usually mammalian, remains. In several cases, the mammalian remains from a set of modern owl pellets collected in the 1990s are used to illustrate an analytical procedure or a concept (see Lyman and Lyman [2003] and Lyman et al. [2001, 2003] for more details on this collection). In the chapters that follow, many points are illustrated by analyzing faunas from various places and dating to various time periods. This helps to emphasize that many properties of the paleozoological record are in at least one sense *universal*, by which is meant that those properties are typically found in an *average* paleo-faunal assemblage. By "average" is meant *typical* and having multiple taxa (usually more than a half-dozen) and multiple identified specimens for the total collection (usually more than, say, 50 specimens). What are rather atypical if not rare or unusual are those collections that have hundreds if not thousands of specimens, all representing the same species. The well-known bison (*Bison* spp.) kill sites of North America do not seem particularly rare because publications on them are numerous, but in terms of the faunal record they are rather atypical. An even more unusual paleofauna would be one consisting of only a couple identified specimens (say, < 10), each representing a unique taxon. When necessary, these sorts of relatively unusual collections are mentioned, but otherwise typical faunas are used to illustrate quantitative concepts and analyses.

Using the same faunal samples throughout, it will be easy to track different kinds of interdependence, and how one analytical result influences whether or not another analysis is reasonable or even feasible. And, by the same token, if two faunal samples from basically the same geographic area and dating to the same time period are available, then other sorts of analytical insights can be gained. Thus two collections of mammal remains are used to illustrate significant points in later chapters. Analyses of the artifacts and features at each site are ongoing, so some of the background information is terse and incomplete. The lack of information on particular aspects of the collections will not make a difference to the points made in this volume.

The collections are zooarchaeological – they originate in an archaeological context. These are faunal remains that had associated artifacts; such assemblages are sometimes referred to as *archaeofaunas*. Both collections consist of mammal remains recovered from two late-prehistoric sites within 10 km of one another. Both sites are found in what is locally known as the Portland Basin or the Wapato Valley of northwestern Oregon state and southwestern Washington state. All mammalian remains from both sites were recovered from one-quarter-inch mesh screens in the field.

The Meier site (35CO5) is located downstream (north) of modern Portland, Oregon, on the floodplain of the Columbia River, on the Oregon side. The Meier site comprises a single large cedar-plank house that was occupied more or less continuously between approximately AD 1400 and AD 1800, and associated midden deposits (Ames 1996; Ames et al. 1992). The site was tested in 1973 and 1984. It underwent extensive excavations every year between 1987 and 1991, inclusively. The 1973 collection was made by Pettigrew (1981) and studied by Saleeby (1983). The 1984 test was directed by Ellis (n.d.); recovered faunal remains have not been analyzed. Kenneth Ames of Portland State University directed the excavations that took place in the late 1980s and early 1990s. I identified all mammalian remains collected by Ames during a 1993 research-sponsored leave when I worked with him in Portland.

The other site, Cathlapotle (45CL1), is located northeast of Meier, on the Washington side of the Columbia River, on a series of levees next to the river. The site was visited by Meriwether Lewis and William Clark in March of 1806 as they lead the Corps of Discovery eastward. At the time of their visit the site comprised several large cedar-plank houses and associated midden deposits (Ames et al. 1999; Ames and Maschner 1999:110). Radiocarbon dates indicate the main occupation began about AD 1450. Ceramic trade goods indicate that abandonment of the site occurred about AD 1834. Auger sampling of the Cathlapotle sediments took place in 1992–1993, and excavations took place each year from 1993 through 1996. Both the auguring and the excavations were under the direction of Ames. I identified all mammalian remains recovered from this site at the University of Missouri-Columbia campus.

The assemblages of mammalian remains from Meier and Cathlapotle were recovered from similar depositional contexts. At both sites, the deposits variously comprise exterior (midden and "yard") deposits and interior deposits (inside of a house). Exterior deposits had very high organic content, lenses of fresh-water mussel shells, and other indications that they formed as primary or secondary dumps (Ames et al. 1999). Yard deposits are generally broad, sheet-like deposits that contain intact hearths, activity areas, pits, evidence of small structures, and so forth. They usually lack the very high organic content of middens though they can have organic content. Interior deposits were assigned to walls, benches (deposits below the 2 m-wide sleeping benches or platforms that ran along the interior side of the house walls), storage pits, and hearth areas. Faunal remains have not been sorted into these distinct depositional contexts as yet. Were they to be so assigned, it is likely that assemblages would be quite small. In later chapters, the influences of small sample sizes receive considerable attention.

The houses at Meier and Cathlapotle had extensive subfloor storage features that, at Meier at least, formed a cellar almost 2 m deep that extended under the house

Table 1.3. *Description of the mammalian faunal record at Meier and at Cathlapotle*

			Cathlapotle			
	Meier		Precontact		Postcontact	
Taxon	NISP	MNI	NISP	MNI	NISP	MNI
*Didelphis**	–	–	–	–	10	1
Scapanus	14	4	–	–	3	2
Sorex	–	–	3	1	1	1
Sylvilagus	16	2	–	–	1*	1
Lepus	–	–	3	2	40	4
Aplodontia	5	1	61	8	57	8
Tamias	1	1	–	–	–	–
Tamiasciurus	2	1	–	–	–	–
Thomomys	9	5	–	–	–	–
Castor	329	9	111	5	238	7
Peromyscus	35	21	2	1	3	2
*Rattus**	1	1	–	–	–	–
Neotoma	1	1	–	–	–	–
Microtus	100	41	10	4	55	22
Ondatra	337	13	56	7	36	4
Erethizon	1	1	–	–	–	–
Canis	90	7	18	3	17	2
Vulpes	2	1	4	1	–	–
Ursus	82	5	45	3	53	4
Procyon	272	20	109	6	84	8
Martes	19	4	1	1	1	1
Mustela	130	17	12	3	9	2
Mephitis	4	2	3	1	–	–
Lutra	45	3	28	3	26	4
Puma	9	1	5	1	5	2
Lynx	22	3	8	1	15	2
Phoca	40	3	26	2	34	2
Ovis	–	–	1	1	1	1
Cervus	832	10	1,091	12	1,793	24
Odocoileus	3,504	56	775	14	1,347	27
*Equus**	–	–	–	–	4	1
TOTAL	5,939	–	2,372	–	3,834	–

*Taxa are historically introduced and not native to the area, hence they are intrusive to site sediments.

floor between the sleeping platforms and the row of hearths in the house's center. The Cathlapotle features are less extensive, but are about 2 m wide by 2 m deep. They are below the sleeping platforms rather than next to them as at Meier. The mammalian remains from both sites were derived primarily from these storage pits and exterior areas.

Because, with few exceptions, the mammalian genera identified are monotypic (include only one species), the basic faunal identification and quantitative data are presented by genus (Table 1.3). Stratigraphy at Meier is extremely complex as a result of multiple episodes of house remodeling and rebuilding; temporally distinct assemblages could not be distinguished. This assemblage is, therefore, treated as a whole and not subdivided into subassemblages. The stratigraphy at Cathlapotle, on the other hand, though also rather complex as a result of the various things that people did at the site, was sufficiently clear that two temporally distinct (sub)assemblages of faunal remains could be distinguished. Many, but not all of the mammal remains from Cathlapotle could be sorted into a pre (Euroamerican) contact sample deposited between AD 1400 and AD 1792, and a postcontact sample deposited between AD 1792 and AD 1835. These two temporally distinct assemblages are referred to in later chapters simply as the precontact and postcontact assemblages. For most purposes the faunal remains that could not be assigned to a temporal period are not included in the analyses presented in later chapters.

The samples of identified remains from the two sites are not tremendously large by some standards (Table 1.3). However, based on data compiled by others (especially Casteel 1977, n.d.), both collections are of reasonable size. That they are not tremendously large is a benefit in the sense that this will assist with the detection of possible influences of sample size in later chapters. The three assemblages are described in Table 1.3. For additional information on the mammal collections not covered in later chapters, see Lyman (2004a, 2006b, 2006c), Lyman et al. (2002), Lyman and Ames (2004), and Lyman and Zehr (2003).

2

Estimating Taxonomic Abundances: NISP and MNI

As paleontologists, Chester Stock and Hildegard Howard were interested in the abundance of the mammals that had walked the landscape and birds that had flown in the air above the landscape at the time the faunal remains from Rancho La Brea were deposited (Chapter 1). Zooarchaeologists (known as archaeozoologists in Europe), on the other hand, are typically interested in which taxa provided the most economic resources and which taxa provided little in the way of economic resources. Thus, as zooarchaeologist Dexter Perkins (1973:369) noted, "the primary objective of faunal analysis of material from an archaeological site [or from a paleontological site] is to establish the relative frequency of each species." This target variable sought by paleontologists and zooarchaeologists concerns *taxonomic abundances*. What are the frequencies of the taxa in a collection?

In any given collection of paleozoological remains, one might wish to know if carnivores are less abundant than herbivores, just as they normally are on the landscape. Given what he knew about ecological trophic structure – that herbivores should outnumber carnivores – imagine Stock's surprise to learn that the typically observed food pyramid or ecological trophic structure was upside down. The mammalian remains from Rancho La Brea represented more carnivores than herbivores – for a reason that many paleontologists thought was a taphonomic reason – because scavenging carnivores got "bogged down" or mired in the sticky tar seeping from the ground and failed to escape. Carnivores were abundant in that tar because they had died and become entombed there as a result of trying to exploit the carcasses of herbivores (and perhaps carnivore brethren) that had themselves become mired and subsequently died there.

The eminently sensible hypothesis that carnivore remains are more abundant than herbivores for taphonomic reasons (see Spencer et al. [2003] for a recent evaluation of this hypothesis) concerns the relationship between a target variable and a measured variable. Stock's target variable was the frequencies of mammalian taxa comprising

the animal community on the landscape, but his measured variable was the sample of bones and teeth from the excavations of the tar pits. Taphonomists have developed an unwieldy terminology (see the glossary in Lyman 1994c), but several of the terms are useful here for keeping *target variables* and *measured variables* distinct. A *biocoenose* is a living community of organisms. Exactly what a community comprises is the subject of some debate. One definition provided by biologists is this: A *community* is comprised of "species that live together in the same place. The member species can be defined either taxonomically or on the basis of more functional ecological criteria, such as life form or diet" (Brown and Lomolino 1998:96).

But "most so-called communities are arbitrary and convenient segments of a continuum of species with overlapping ecological requirements, not involving a high level of interdependence" (Lawson 1999:7). Thus one commentator notes that a biological community can be defined one of two ways: As a group of organisms occupying a location, or as a group of organisms with ecological linkages among them (Southwood 1987). Often the focus is on the former at the expense of the latter; a community (which may include all organisms or a particular subset of organisms) is often defined by specific spatial boundaries (Magurran 1988:57). Perhaps not surprisingly, the "nature of boundaries separating adjacent communities is hotly disputed by community ecologists and paleoecologists" (Hoffman 1979:364). For the sake of discussion, we assume that a biological community and its boundaries can be defined on the landscape today.

The taxonomic composition and taxonomic abundances of a biocoenose might well be the target variable sought by a paleozoologist. However, that biocoenose is not what a paleozoologist studies. The organisms that comprise a biocoenose must die before a paleozoologist can study their mortal remains. A *thanatocoenose* is an assemblage of dead organisms; it is sometimes referred to as the death assemblage. Turner (1983) suggested that it be referred to as the "killed population," but "killed" implies an active agent of death, such as a predator, disallowing death from other causes such as old age. Furthermore, the dead organisms may be a "population" in a statistical sampling sense, but they might not be, depending on the question asked. Those dead organisms may be a 100 percent sample of some set of dead organisms (e.g., all of those from a biocoenose, which, after all, must all die), but they may be less than that, such as when the set of dead organisms represents only part – a sample – of a biocoenose. In this case, the thanatocoenose is not, statistically speaking, a "population." It is perhaps best, for this reason alone, to conceive of a thanatocoenose as a set of dead organisms, usually somehow stratigraphically or analytically bounded (see the discussion of a *faunule* later in this chapter).

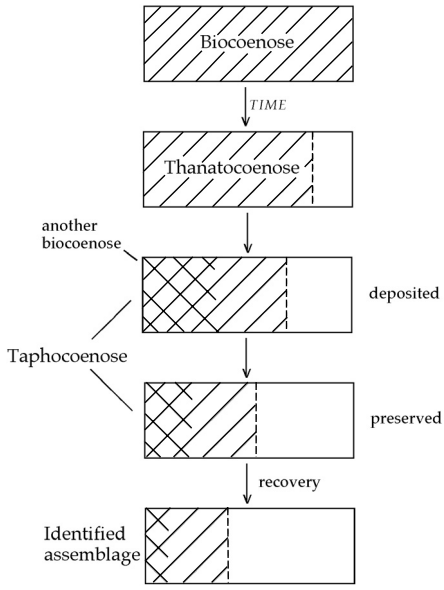


FIGURE 2.1. Schematic illustration of loss and addition to a set of faunal remains studied by a paleozoologist.

Given that paleozoologists sample the geological record (i.e., where faunal remains are deposited as a particular kind of sedimentary particle), they don't always have a complete thanatocoenose lying on the lab table. Furthermore, the organisms whose remains comprise a thanatocoenose may derive from one community (or biocoenose) or they may derive from more than one community (Shotwell 1955, 1958). A *taphocoenose* is the set of remains of organisms (in our case, faunal remains) found buried (or perhaps exposed) and spatially (usually stratigraphically) associated. Given that not all of the remains comprising a taphocoenose will be recovered, and of those that are recovered not all will be identified to taxon, what the paleozoologist can identify comprises what will here be referred to as the *identified assemblage*. It is this set of remains from which measures of taxonomic abundance are derived.

Figure 2.1 is a schematic rendition of the typical differences between a biocoenose and the identified assemblage. A biocoenose is a biological community. One or more biocoenoses is the source of input to a thanatocoenose. The transition from biocoenose to thanatocoenose involves *accumulation* (and deposition) of faunal remains in a location; accumulation can be *active* (involve a bone-accumulating agent, such as a predator that transports prey to a den) or *passive* (involve deaths of animals across the landscape, referred to as "background accumulation" [Badgley 1986]).

The transition from the thanatocoenose to taphocoenose involves both accumulation and *dispersal* and movement or removal of bones mechanically (such as by fluvial transport), and also the deterioration or chemical and mechanical breakdown of skeletal tissue. The transition from the taphocoenose to the identified assemblage involves both recovery (usually < 100 percent for any of several reasons) by the paleozoologist and the taxonomic identification of a subset of the remains comprising the taphocoenose.

The measured variable of taxonomic abundances originates in the identified assemblage; the target variable depends on the question asked (Figure 2.1). The biocoenose was Stock's target variable. Zooarchaeologists interested in human subsistence and economy often have, as their target variable, a thanatocoenose created by human predators. Paleoecologists, whether working as paleontologists or as zooarchaeologists who are interested in past ecological conditions, have as a target the biocoenose. Determination of the statistical relationship between an identified assemblage and a target variable is a taphonomic concern (Lyman 1994c), but the paleozoologist should also consider ecological and animal behavior variables as well as recovery techniques. As the Rancho La Brea materials indicate, animal behavior can influence the accumulation and thus rate of input of animal remains to the geological record.

Not all taxonomic abundances at Rancho La Brea can be attributed to historical contingencies of how and why the faunal remains present were originally accumulated. Did, perhaps, local habitats support more artiodactyls than perissodactyls, or as with the carnivores at Rancho La Brea, did a particular bone-accumulation agent bring more even-toed herbivores than odd-toed herbivores to the tar pits? Whatever the case, we measure taxonomic abundances in the identified assemblage and use those values as proxy measures or estimates of a thanatocoenose or a biocoenose. It is the task of taphonomic analyses to ascertain how good (or how biased) an estimate of a particular target variable the measured variable might be. Paleozoologists have, over the past 20 years or so become very concerned over how good an estimate of a biocoenose an identified assemblage might provide (e.g., Gilbert and Singer 1982; Ringrose 1993; Turner 1983). This concern has resulted in research known as *fidelity studies*, or assessment of "the quantitative faithfulness of the [fossil] record of morphs, age classes, species richness, species abundance, trophic structure, etc. to the original biological signals" (Behrensmeyer et al. 2000:120).

Comparisons of faunal remains with organisms making up a biological community from which the remains derive suggest that fidelity can range from quite high to rather low with respect to the variables of interest, and not all variables display equivalent fidelity in any given set of remains (Hadly 1999; Kidwell 2001, 2002; Kowalewski et al.

2003; Lyman and Lyman 2003). If there is no taphonomic reason for artiodactyl remains to outnumber perissodactyl remains at Rancho La Brea, for example, then an ecological explanation – there were more artiodactyls than perissodactyls on the landscape to be accumulated because climatic conditions created habitats more favorable to artiodactyls – is likely.

The distinction between measured taxonomic abundances and target taxonomic abundances will be important throughout the remainder of this chapter, so keep it in mind. Turner (1983) suggests that often one must assume that the relative taxonomic abundances evident in an identified assemblage are a statistically accurate reflection of those abundances in a taphocoenose, a thanatocoenose, and a biocoenose. This is true, but it is also incomplete in the sense that we can do more than assume accurate reflections; we can often test the general accuracy of the reflection with data independent of the identified assemblage at hand. The fauna represented by the identified assemblage should align ecologically with, say, floral (pollen, phytoliths, plant macrofossils) data. If it does not, then either the identified faunal assemblage is not an accurate reflection of the biocoenose, or the plant record is not. Similarly, an identified faunal assemblage at a nearby site (assuming similar ages) should align with the assemblage under consideration; two taphonomically independent assemblages should have statistically indistinguishable taxonomic abundances (and the more taphonomically independent samples that indicate the same biocoenose, the more accurate the conclusion). If they do not, one or both of the assemblages may not be a good reflection or estimate of taxonomic abundances within the local biocoenose (Grayson 1981a; Lundelius 1964).

Regardless of whether one is interested in the taphonomic history of a collection of faunal remains (how and why those remains were differentially accumulated, deposited, preserved [and some would say recovered]), in the biogeographical implications of the taxa represented by those remains (why are these taxa here but not other taxa), in the paleoecological implications of the represented taxa (do the represented species signify warm or cool climates, or moist or dry habitats), or in the subsistence and foraging behaviors of the accumulation organism (human if an archaeological site, a carnivore if a den), *taxonomic abundances* are typically part of the data scrutinized for answers to the research question. As Grayson (1979:200) noted, "It is virtually impossible to find any faunal analysis that does not present one or more measures of taxonomic abundance. [This variable] is a basic one." The critical tactical decision concerns choosing a method to measure the abundances of the taxa represented in a collection. It is easy to show that this is no simple matter.

When Stock tallied up the remains from Rancho La Brea (Figure 1.1), it is likely that he reflected on the fact that proboscideans have more bones in one skeleton

than do perissodactyls. The former have five digits at the end of each limb whereas various perissodactyls, late Pleistocene equids in particular, have only one digit at the end of each limb. Thus, simply tallying up how many bones (and teeth) each taxon contributed to a collection could potentially produce inaccurate estimates of the abundances of the various taxa represented. If each skeleton of taxon A produces 75 identifiable bones, and each skeleton of taxon B produces 90 identifiable bones, then if 5 individuals of each were mired at Rancho La Brea, one would have 5 skulls of each taxon but 375 bones (number of identified specimens, or NISP) of taxon A and 450 bones (= NISP) of taxon B.

Ignoring for the moment potential differences in the number of teeth various taxa may have, simple tallies of the skeletal specimens of taxa A and B could produce misleading insights to which one of the two was more abundant on the landscape. Note that NISP is the measured variable whereas taxonomic abundances in the biocoenose is the target variable. Recognizing the differences between these two variables is critical to understanding whether a measure is valid or not. Is the target variable, for example, the frequency of each taxon recovered from the site (identified assemblage), the frequency of each taxon preserved in the site sediments (taphocoenose), the frequency of each taxon accumulated and deposited in the site (taphocoenose, thanatocoenose), or the frequency of each taxon available on the landscape (biocoenose)? This question underscores the simple fact that accumulation, deposition, preservation, recovery, and identification of faunal remains can all weaken in unpredictable ways the statistical relationship between the measured variable and the target variable.

If a measure of taxonomic abundances is desired, then what sort of quantitative unit should be used? Obviously, we want a unit that allows us to estimate taxonomic abundances in a sample of bones, teeth, and shell lying on the lab table. That is, we want a unit that measures taxonomic abundances within the identified assemblage; how closely those abundances match up with taxonomic abundances in the thanatocoenose or biocoenose from which the identified assemblage derives is a separate question that requires detailed taphonomic analyses and other sorts of data. We cannot presume that taxonomic abundances are the same across the identified, taphocoenose, thanatocoenose, and biocoenose assemblages. But as noted above, we can sometimes perform empirical tests to determine if this is so or not. In this chapter the two fundamental quantitative units originally designed to measure taxonomic abundances are discussed. These quantitative units are known as NISP and MNI; they are considered in turn. Biomass and meat weight and other quantitative methods used to measure taxonomic abundances are discussed in Chapter 3.

THE NUMBER OF IDENTIFIED SPECIMENS (NISP)

The most fundamental unit by which faunal remains are tallied is the number of identified specimens, or NISP. It is just what it sounds like – the number of skeletal elements (bones and teeth) and fragments thereof – all specimens – identified as to the taxon they represent. A related measure sometimes mentioned is the number of specimens (NSP) comprising a collection or assemblage. The NSP includes bones, teeth, and fragments thereof some of which have been identified to taxon, plus those specimens that have not or cannot be identified to taxon. Typically, "identified to taxon" means identified as to the skeletal element and to the taxonomic order, family, genus, or species represented by the specimen. Most taxonomically diagnostic anatomical features are also diagnostic of skeletal element (is it a humerus or a tibia?). Many paleozoologists do not tally nondescript pieces of bone that are from the taxonomic class Mammalia if those pieces cannot be assigned to taxonomic order, family, genus, or species. This is so because taxonomic identifications such as "mammal long bone fragment" generally, but not always, are of little analytical utility. But don't misunderstand. The research problem or question one is grappling with should, if carefully phrased, indicate whether or not an otherwise nondescript piece of "mammal long bone" is worthy of tallying or not. For the remainder of this book, "identified" means that a specimen has been minimally identified as to skeletal element and to at least taxonomic family (if not genus or species), unless otherwise noted. As Stock's (1929) example summarized in Chapter 1 makes clear, much may be learned from taxonomic order-level identifications.

Virtually all paleozoological collections consist of some NSP of which a fraction makes up the NISP. NISP is the number of identified (to skeletal element and at least taxonomic family) specimens determined for each taxon for each assemblage. When one says NISP, what is meant is NISPi where i signifies a particular taxon. This is analogous to statistical symbolism because the i is seldom shown; rather, it is understood. Thus, one has an NISP of 10 for deer (*Odocoileus* sp.) and an NISP of 5 for rabbits (*Sylvilagus* sp.). In some cases the symbolism may be more complex, such as NISPij, where i is for the taxon as before, and j is for a particular skeletal element or part, such as a humerus or distal tibia. Again, the ij is not shown but is understood. It is likely for these reasons of implied symbolism that it is unusual to see the plural form of NISP, such as NISPs; in my experience (which may not be representative), it is more common to read "NISP values" when more than one taxon or skeletal element is intended, as when an NISP value is given for each of multiple taxa.

Advantages of NISP

The acronym NISP and its meanings, both implied (ij) and explicit (number of identified specimens), should be clear. Its operationalization may also seem to be clear and straightforward. For any pile of faunal remains, identify every specimen that you can (to skeletal element and taxon), and then tally up how many specimens you identified per taxon (and perhaps also noting how many specimens represent each kind of skeletal element, depending on your research question). NISP is an observed measure because it is a direct tally, and so it is not subject to some of the problems that derived measures such as MNI are. In part because NISP is an observed or fundamental measure, it has advantages over other units used to measure taxonomic abundances. First, NISP can be tallied as identifications are done. That is, NISP is additive or cumulative; the analyst does not have to recalculate NISP every time a new bag of faunal remains is opened and new specimens identified. This property makes NISP a fundamental measure. Every identified specimen represents a tally of "1." Add up all the tallies of "1" for each taxon to derive the total NISP per taxon or $\sum$NISPi.

NISP is, however, not free of problems. One long recognized (Clason 1972; Lawrence 1973) but seldom mentioned problem influences both NISP and MNI. Different analysts may identify different specimens in a pile of faunal remains (Gobalet 2001). The sets of specimens that any two analysts identify will be quite similar – all complete teeth and complete limb bones are likely to be identified, assuming both analysts have access to similar comparative collections – but they may not be identical, which means interanalyst comparability is imperfect. Whether a particular specimen is identifiable or not depends on the anatomical landmarks available on that specimen (Lyman 2005a), and experience and training will influence what an individual analyst will identify because that experience and training dictates which landmarks the paleozoologist has learned are useful.

Interobserver difference in what is identified can be a serious source of variation in NISP tallies. Because it concerns what is identified, interobserver difference applies to any conceivable measure of taxonomic abundance. As with other kinds of interobserver difference, it is not just difficult to control. It is impossible to control (unless every paleozoologist studies every collection) and it is likely for this reason that few paleozoologists have mentioned it. It is mentioned here for sake of completeness and because it may be an important consideration when one analyst compares his or her tallies with someone else's for a different collection. Because it cannot be controlled, it is not considered further. But there are other potential problems with NISP. One might think that interobserver variation in how to tally what is identified for

purposes of producing an NISP value will not vary from investigator to investigator because each identified specimen represents a tally of 1. But, does it? Answering this question brings us to some of the weaknesses internal to NISP, weaknesses identified and described by many researchers.

Problems with NISP

When Stock (1929) presented his census of Rancho La Brea mammals, he tallied the minimum number of individual(s) animals – what are now called MNI values – rather than NISP. It is likely he did so because he recognized that members of the Carnivora have different numbers of first (or proximal) phalanges per individual (usually 4 or 5 per limb) than do the Perissodactyla (Pleistocene horses have one) or the Artiodactyla (usually two). NISP tallies of first phalanges for a single dog would be 16 (ignoring the vestigial first + second phalanx of the first digit of each foot), for a single horse the NISP of first phalanges would be 4, and for a cervid the NISP of first phalanges would be 8.

Many problems with using NISP to measure taxonomic abundances have been described over the years by numerous authors (e.g., Bökönyi 1970; Breitburg 1991; Chaplin 1971; Gautier 1984; Grayson 1973, 1979; Higham 1968; Hudson 1990; O'Connor 2001, 2003; Payne 1972; Perkins 1973; Ringrose 1993; Shotwell 1955; Uerpmann 1973). Long lists of the possible weaknesses and potential problems with using NISP as a measure of taxonomic abundances are given by Grayson (1979, 1984). The following is based on his lists, and is supplemented with concerns expressed by paleontologists (e.g., Shotwell 1955, 1958; Van Valen 1964; Vermeij and Herbert 2004):

1 NISP varies intertaxonomically because different taxa have different frequencies of bones and teeth (the number of elements that are identifiable varies intertaxonomically);
2 NISP will vary with variation in fertility (number of offspring per reproductive event) and fecundity (number of reproductive events per unit of time);
3 NISP is affected by differential recovery or collection (large specimens [of large organisms] will be preferentially recovered relative to small specimens [generally of small organisms]);
4 NISP is affected by butchering patterns (different taxa are differentially butchered, one result of which is intertaxonomic differential accumulation of skeletal parts, and another of which is intertaxonomic differential fragmentation of skeletal elements);

5 NISP is affected by differential preservation (similar to problem 4; taphonomic influences may vary intertaxonomically);
6 NISP is a poor measure of diet (the bones of one elephant provide more meat than the bones of one mouse);
7 NISP does not contend with articulated elements (is each tooth in a mandible tallied as an individual specimen, plus the mandible itself tallied?);
8 the problems identified may vary between strata within a site, between distinct sites, or both, rendering statistical comparison of site or stratum specific assemblages invalid;
9 NISP may differentially exaggerate sample sizes across taxa;
10 NISP may be an ordinal scale measure and if so some powerful statistical analyses are precluded as are some kinds of inferences;
11 NISP suffers from the potential interdependence of skeletal remains.

Problems, Schmoblems

The list of problems analysts have identified as plaguing NISP values may seem disconcerting. Indeed, the length of the list may give one cause to wonder why anyone would measure taxonomic abundances using NISP in the first place. Do not, however, let such wonder convince you that NISP values are worthless. Some problems overlap with one another in terms of their effects or in terms of how they might be dealt with analytically. And notice that the list comprises a set of "possible weaknesses and potential problems." Several of the problems are easily dealt with analytically.

Problem 1 can be controlled in several ways, such as only counting elements held in identical frequencies by the taxa under study (e.g., Plug and Sampson 1996). Do not tally phalanges of artiodactyls and perissodactyls when comparing their abundances; tally only scapulae, humeri, femora, and other elements that occur in identical frequencies in individuals of both taxa. In short, do not include tallies of elements that vary in frequency intertaxonomically. Or, weight NISP by dividing it by the number of identifiable elements per single complete skeleton in each taxon. So, if, say, horses always have 100 elements per complete skeleton and bison always have 85 elements per skeleton, then weight observed abundances of NISP for horses and for bison accordingly. This solution was suggested more than 50 years ago by paleontologist J. Arnold Shotwell (1955, 1958). It is, however, not without problems, such as requiring the assumption that complete skeletons (rather than a limb or two) were accumulated and deposited in the collection location. The assumption comprises a taphonomic problem, and might be addressed by consideration of which

skeletal elements are present. What about variation in rates of input of skeletal parts to the geological record?

We don't need to worry about correcting for differences in number of skeletal elements per taxon because, to retain the example, late-Pleistocene horses always have the same number of skeletal elements in each of their skeletons as every other late-Pleistocene horse, and late-Pleistocene bison have the same number of skeletal elements in each of their skeletons. Thus, we know that if the NISP of bison increases relative to the NISP of horses (the measured variables), then perhaps the abundance of bison (on the landscape, or in the identified assemblage) increased relative to the abundance of horses (the target variables). Bison NISP did not increase relative to horse NISP because bison evolved to have more bones or horses evolved to have fewer bones over the time represented. The same argument applies to variables that influence the rate of input of skeletal parts to the faunal record. Shotwell (1955, 1958) was concerned that different taxa input bones to the paleozoological record at different rates. A few years later Van Valen (1964) spelled out his concern that different taxa have different longevities; taxa with short individual life spans input more skeletal parts to the faunal record than taxa with long individual life spans, all else being equal (same number of skeletal parts per taxon, same population size on the landscape).

Problem 2 was recently stated by Vermeij and Herbert (2004), who worried that intertaxonomic variation in fertility and fecundity influenced the rate of skeletal part input. They noted that "short-lived (often small-bodied) species will be greatly over represented in a fossil sample relative to species with long generation times, long individual life spans, slow rates of turnover, and large body size" (Vermeij and Herbert 2004:2). If taxon A has an individual average life span of 10 years whereas taxon B has an individual average life span of 1 year, then taxon B will be represented by ten times the number of individuals as taxon A (all else being equal). Vermeij and Herbert (2004:3) were concerned that measures of "predator-prey ratios and prey availability" would be artifacts of variation in life span. Their primary solution to problem 2 requires data on average life spans, in some cases derivable from the growth increments evident in the hard parts of organisms. In the absence of the requisite ontogenetic data, a secondary solution they suggest is to restrict sampling "to organisms of comparable generation time," though this solution also seems to require taxon-specific ontogenetic information.

The first solution is, in fact, identical in reasoning to the one suggested by Shotwell (1955, 1958) for the problem of mammalian taxa with different numbers of (identifiable) skeletal elements per individual. They are "identical" because both Shotwell and Vermeij and Herbert were concerned about biological factors that influence the

Table 2.1. *Fictional data on abundances (NISP) of three taxa in five strata*

Stratum	Taxon A	Taxon B	Taxon C	Total
V	50 (77)*	10 (15)	5 (8)	65
IV	40 (67)	10 (16.5)	10 (16.5)	60
III	30 (55)	10 (18)	15 (27)	55
II	20 (40)	10 (20)	20 (40)	50
I	10 (22)	10 (22)	25 (56)	45

* Relative (percentage) abundances of each taxon given in parentheses.

rate at which skeletal remains are created and input to the paleozoological record. Taxa with many skeletal elements per individual and taxa with high fecundity both have higher rates of input than taxa with few skeletal elements per individual and taxa with low fecundity, respectively. When faced with the former problem, Shotwell suggested that the analyst should determine a "corrected number of specimens" per taxon, a value calculated by dividing each taxon's NISP by the number of identifiable elements in one skeleton of that taxon. If an individual skeleton of taxon A potentially contributes 10 (identifiable) elements and taxon B 5 elements, then divide the observed NISP for A by 10 and the observed NISP for B by 5 in order to compare the abundances of the two taxa. A similar procedure for invertebrates is described by Kowalewski et al. (2003). The procedures norm all taxon-specific NISP values to a common scale – the number of identifiable skeletal elements per individual per taxon. In light of Vermeij and Herbert's concern, a paleozoologist could norm all taxonomic abundances to a single life span, based on the duration of all life spans measured in the same unit, say, a year.

The concerns of Shotwell, Van Valen, and Vermeij and Herbert are all easily dispensed with. Table 2.1 lists fictional NISP values for three taxa in five strata. Because we know that taxon A has 10 skeletal elements per individual, taxon B has 1 skeletal element per individual, and taxon C has 5 skeletal elements per individual, we choose to weight their abundances accordingly. The results of that weighting are shown in Table 2.2. For ease of conceptualizing what has happened, consult Figure 2.2. Note that over the five strata, whether the NISP values are the raw tallies or the weighted tallies corrected for differences in number of skeletal elements per taxon, taxon A increases from stratum I to stratum V whereas taxon C decreases over that same span. Weighting does not change the results, at least with respect to increases and decreases in the *relative* abundances of taxa A and C. But, you might counter, in the unweighted data taxon B is often not very abundant at all, and it too gradually decreases from stratum I to stratum V. In the weighted data, however, taxon B is more abundant than

Table 2.2. *Data in Table 2.1 adjusted as if each individual of taxon A had ten skeletal elements per individual, taxon B had one skeletal element per individual, and taxon C had five skeletal elements per individual*

Stratum	Taxon A	Taxon B	Taxon C	Total
V	5 (31.5)*	10 (62.5)	1 (6.25)	16
IV	4 (25)	10 (62.5)	2 (12.5)	16
III	3 (18.75)	10 (62.5)	3 (18.75)	16
II	2 (12.5)	10 (62.5)	4 (25)	16
I	1 (6.25)	10 (62.5)	5 (31.25)	16

* Relative (percentage) abundance in parentheses.

A and C combined, and taxon B doesn't change in relative abundance over the stratigraphic sequence. That is a good point – it suggests the data are ordinal scale – and we will return to it. First, however, we need to consider other problems with NISP.

Problem 3 concerns collection bias. Correction factors might be designed to account for the fact that small bones and teeth and shells tend to come from small organisms, and these tend to escape visual detection and to fall through coarse-meshed hardware cloth meant to allow the passage of sediment but not faunal

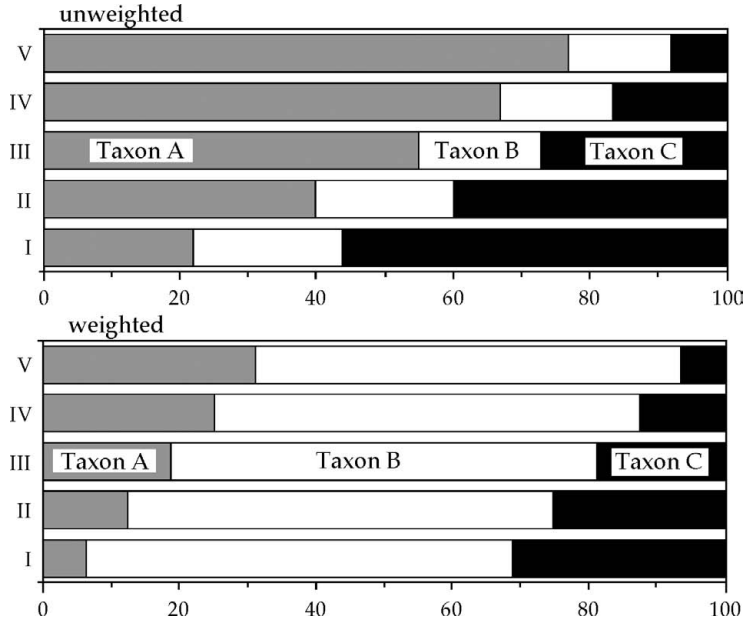


FIGURE 2.2. Taxonomic relative abundances across five strata. Data from Tables 2.1 (unweighted) and 2.2 (weighted).

remains. The design of correction values has also been a long-standing interest among zooarchaeologists (e.g., Payne 1972; Thomas 1969), but again, there are problems with these values. For example, if one uses correction values, one must assume that the samples used to derive those values are on average representative of all situations – within any given excavation unit, within any given stratum of a site, and within any given site – where small remains may fall through screens (see Chapter 4). As long as recovery methods do not differ between strata, such as using 1/4-inch mesh hardware cloth for every other stratum and using 1/8-inch mesh hardware cloth for the other strata, remains of mice will be as consistently recovered in all recovery contexts as are remains of rabbits and deer.

The preceding does not allow for differential fragmentation across taxa, the issue raised by problem 4. If rabbit bones are quite fragmented and small, they may fall through screens much more readily than unfractured remains of mice (Cannon 1999). Large bones may be more likely to be fractured by humans or carnivores because they contain more nutrients (marrow) than small bones. Fragmentation reduces identifiability by disassociating if not destroying the distinctive landmarks used to tell that bone A is a rabbit tibia whereas bone B is a duck humerus (Lyman 2005a; Marshall and Pilgram 1993). Problem 5 concerns intertaxonomic differences in preservation and may be related to problem 4 given that fragmentation influences preservation by effectively destroying bones through the process of rendering them unidentifiable. If preservational processes vary intertaxonomically, then NISP values will be differentially influenced across taxa. The magnitude of the fragmentation problem can be evaluated analytically (see Chapter 6). Fragmentation and preservation are taphonomic processes and may well render NISP data nominal scale with respect to taxonomic abundances.

Problem 6 is a serious concern to many zooarchaeologists because it is true that NISP is a poor measure of diet because the meat from the bones of one elephant will feed more people than the meat from the bones of one mouse. Furthermore, ethnoarchaeological data suggest that we cannot assume that each individual animal carcass was consumed entirely (e.g., Binford 1978; Gifford-Gonzalez 1989; Lyman 1979). But if we are concerned with dietary issues, we have in fact changed the target variable from a measure of taxonomic abundances – are there more elephants represented in the collection, or more mice – to one concerning how much of each taxon was eaten. Because we are asking a different question, a quantitative unit or measured variable different than one that simply tallies taxonomic abundances would seem to be required (see Chapter 3).

Problem 7 is that NISP does not include inherent rules for dealing with articulated elements, and while rules have been suggested (e.g., Clason 1972), these are not agreed upon by all paleozoologists or used consistently. A common example concerns

what to do with a mandible or maxilla that contains teeth. The mandible (dentary) is a discrete anatomical organ, as is each tooth. In ungulates that means a single mandible (say, the left side) containing all teeth will have 4 incisforms (3 incisors and a canine that has evolved into the form of an incisor), 3 deciduous premolars, 3 permanent premolars, and 3 molars. It is rare to find mandibles with all 6 premolars (the deciduous ones nearly worn to nothing and about to fall out of the mandible; the permanent ones still forming and without developed roots but beginning to erupt). So, ignoring that possibility for the moment, when a mandible with all teeth is found, do we tally an NISP of 1, or do we tally an NISP of 11 (mandible + 4 incisiforms + 3 premolars + 3 molars)? How do we tally an articulated hind limb of an artiodactyl; as 1, or as 15 (femur + patella + tibia + distal fibula [lateral malleolus] + calcaneus + astragalus + naviculo cuboid + 4th tarsal + metatarsal + 6 phalanges [not to mention sesamoids])?

The paleozoologist should be consistent in applying across all taxa the tallying method chosen, and also be explicit about which method is used – tally articulated specimens as 1, or tally each distinct anatomical element, articulated or not, as 1. Perhaps the most important aspect of dealing with this problem is granting flexibility to meet the needs of analysis and the nature of the collection. Each mandible, for example, can be tallied as 1 regardless of whether it includes teeth or not (noting which teeth if any are present for purposes other than estimating taxonomic abundances, although ontogenetic age differences indicated by the teeth may play a role in estimating abundances [see the discussion of MNI]). But which skeletal elements were articulated when found and which were isolated or not articulated with other specimens is seldom noted by excavators. Thus, the paleozoologist may be forced to tally each specimen individually as 1, noting when one specimen "articulates" with another if they come from the same excavation unit, which is not necessarily the same as saying that it was "articulated" when it was found.

Noting that the problems listed thus far may vary not only intertaxonomically, but intrataxonomically within and between strata (problem 8) sounds hopelessly fatalistic. However, it is largely an analytical matter to determine if indeed this taphonomic problem applies in any given case. Even if it does, it may not preclude statistical comparisons of assemblages of faunal remains. Before arguments and examples of why this is so are presented, we consider what is likely the most serious problem with NISP. As a preface, note that most of the problems with NISP discussed so far are *not* fatal to it as a measure of taxonomic abundance. Most of these problems were identified and subsequently reiterated time and time again not as reasons to abandon NISP and design a new quantitative unit but rather were presented as warrants to use MNI (Grayson 1979, 1984). A prime example of this is problem 6. That problem – that NISP doesn't give a good estimate of the amount of meat provided – is like

Table 2.3. *The differential exaggeration of sample sizes by NISP*

	Taxon A	Taxon B	Taxon C
NISP	1	10	10
MNI	1	1	10

saying that because a tape measure doesn't measure color, the tape measure is flawed. Of course, the tape measure was not designed to measure color, or weight, or material type; rather, it was designed to provide measures of linear distance. Because a measurement unit doesn't measure a particular variable is no reason to discard that unit completely. No one has demonstrated that NISP doesn't provide valid and accurate measures of taxonomic abundances in a taphocoenose, in a thanatocoenose, or in a biocoenose. Indeed, virtually all of the problems with NISP do not universally invalidate it as a unit with which to measure taxonomic abundances.

None of the preceding is meant to imply that NISP is a valid measure of taxonomic abundances in a taphocoenose, in a thanatocoenose, or in a biocoenose. It might be a valid measure of taxonomic abundances, but that remains to be determined. Before discussing how to make that determination, the most serious problem with NISP must be identified. This problem must be considered at length precisely because it is so worrisome.

A Problem We Should *Worry About*

That NISP may differentially exaggerate sample sizes across taxa (problem 9) is evident in the example in Table 2.3. This table illustrates that if one has an NISP of 1, then at least 1 individual (MNI) is represented; if NISP = 2, then MNI = 1 or 2; if NISP = 3, then MNI = 1, 2, or 3; and so on. Thus, were we to compare the abundances of taxa A, B, and C in Table 2.3, taxon B would be over-represented by NISP relative to taxa A and C. This is so because for taxa A and C, each individual (MNI) is represented by one specimen, so each specimen contributes an MNI of 1. But for taxon B, each individual is represented by an NISP of 10, so each specimen in effect contributes one-tenth of an individual MNI.

Problem 10 is that NISP is an ordinal-scale measure of abundance so some powerful statistical analyses and inferences are precluded. We can often say that taxon A is more abundant than taxon B, but we do not know by how much with respect to a target variable consisting of the thanatocoenose or the biocoenose. This is so even when we

can control for variation in fragmentation, variation in the number of identifiable elements per individual of different taxa, and all those other problems that afflict NISP. Notice that I said we can "often" say that one taxon is more abundant than another; I did not say that we can "always" say this. We return to this point later in this chapter.

The potential overrepresentation of some taxa by NISP is in fact a superficial concern, but it is intimately related to the deeper, more serious concern expressed in problem 11 – NISP suffers from the fact that skeletal specimens may be interdependent (Grayson 1979, 1984). The specimens of a taxon in a collection, or various subsets of those specimens, may be from the same individual animal (or each subset from a different individual). This precludes various statistical analyses and tests of taxonomic abundance data tallied as NISP that demand independent data, that is, each tally of "1" is independent of every other one. Some analysts have argued that specimen interdependence is not a serious problem. Gautier (1984), for example, based on estimates of preservation rates at various sites, notes the probability that an animal would be represented by a single specimen, by two specimens (the product of two independent probabilities represented by two specimens), by three specimens (the product of three independent probabilities), and so on. He finds that the probabilities for NISP > 1 for any given individual are quite low, so Gautier (1984:240) concludes "the degree of interdependence (i.e., the fact that an animal is represented by several bones and hence counted several times) is much less than many analysts fear."

Gautier's (1984) estimates of preservation rates are based on a compilation of many estimates – the estimated duration of occupation of the site in years, the estimated size of the human population that occupied the site, the estimated number of animals necessary to provide sufficient food for the human occupants, the estimated degree of preservation of faunal remains, the estimated fraction of the site excavated, and the estimated rate of identification of faunal remains. As these estimates are added up, one influencing another, the final estimate of whether two specimens derive from the same animal is likely quite wide of the mark. The estimate is like a radiocarbon age of 1,000 years with an associated standard deviation of 900 years. Furthermore, Gautier's estimates must assume that the taphonomic history of each specimen is independent of the taphonomic history of every other specimen, even when two specimens derive from the same individual animal. We know that that is false (Lyman 1994c), else we would never find articulated bones.

So, presuming that there is some degree of interdependence of specimens tallied for NISP values, what is the paleozoologist to do? One option is to accept Gautier's arguments, and as Perkins (1973:367) suggests, "in the absence of archaeological evidence to the contrary we must assume that each [specimen] came from a different

individual." This allows statistical manipulation of NISP data as if each tally of 1 for each taxon was indeed independent of every other tally of 1 for that taxon. But it is also likely that skeletal elements of individual animals were not accumulated and deposited completely independently of each other (Ringrose 1993). They were, after all, articulated and held together during the life of the organism. Actualistic work indicates that although complete skeletons may not accumulate as such, portions of skeletons comprising multiple elements are very often accumulated by both human and nonhuman taphonomic agents (e.g., Binford 1978, 1981; Blumenschine 1986; Domínguez-Rodrigo 1999a; Haynes 1988; Lyman 1989, 1994c). This brings us back to the question at the beginning of this paragraph: Given that there is some unknown (and largely unknowable) degree of interdependence in NISP values, what is the paleozoologist to do?

One thing we might do is assume that interdependence is randomly distributed across all taxa and all assemblages (Grayson 1979). Assuming this does not, of course, make it so. But if we could show that interdependence was distributed across taxa and assemblages in such a way as to not significantly influence measures of taxonomic abundances, then we would have an empirical warrant for using NISP to measure those abundances. So, the question shifts from: "Given the likelihood that there is some interdependence, what are we to do?" to "How are we to show that the nature and degree of interdependence does not significantly influence NISP measures of taxonomic abundance?" The answer to the new question must come after we consider the other quantitative unit that is regularly used to measure taxonomic abundances.

THE MINIMUM NUMBER OF INDIVIDUALS (MNI)

Given the many difficulties with using NISP to measure taxonomic abundances, it is not surprising that Stock (1929) and Howard (1930) estimated abundances of mammals and birds at Rancho La Brea (Chapter 1) using a measure other than NISP. The measure they used is the *minimum number of individuals* (MNI). Prior to the middle 1990s, a plethora of acronyms were used for this quantitative unit (Lyman 1994a). As with NISP, MNI usually (but not always) is given for each identified taxon, so the acronym is more completely given as MNIi, where i again signifies each distinct taxon and, again, is seldom shown but is instead understood. Recall that Stock and Howard both defined MNI as the most commonly occurring skeletal element of a taxon in an assemblage. Thus, if an assemblage consists of three left and two right scapulae of a species of mammal, then there must be at least (a minimum of) three individuals of that species represented by the five specimens because each

individual mammal has only one left and one right scapula. The number of individuals is a *minimum* because there may actually be five individuals represented by the five scapulae, but it presently is difficult to determine in each and every case which left scapula goes with which right scapula (come from the same individual), nor can we always determine when potentially paired elements do not come from the same individual. Thus, the *actual number of individuals* (ANI) represented by the identified assemblage of a taxon is difficult to determine, except perhaps in those rare cases when the taxon is represented by more or less complete articulated skeletons.

MNI is an attractive quantitative unit because it solves many of the problems that attend NISP. In particular, it solves the critical problem of specimen interdependence given how MNI is usually defined – the most commonly occurring kind of skeletal specimen of a taxon in a collection. This is indeed how many (but certainly not all) analysts define the term (Table 2.4). If the most commonly occurring kind of skeletal specimen of taxon A is distal left tibiae, then tally up distal left tibiae; the total equals the MNI of taxon A. If the most commonly occurring kind of skeletal specimen of taxon B is the right m3, then tally those up and the total gives the MNI for taxon B. No single individual of any known mammalian taxon possesses more than one distal left tibia or more than one right m3, so each one of those kinds of specimens in a collection must represent a unique individual that was alive in the past. An easy way to conceptualize MNI is this: If two skeletal specimens *overlap* anatomically, then they must be from distinct, independent individual organisms because they are redundant with one another (Lyman 1994b). If the two specimens fit together in a manner like two conjoining pieces of a jigsaw puzzle, then they are from the same individual and are interdependent. But if the two specimens do not overlap anatomically and they do not fit together like two pieces of a jigsaw puzzle, then they *may* be from the same individual unless they are clearly of different size, ontogenetic (developmental) age, or sex. If they are of the same size, age, and sex, but do not overlap or conjoin, then they *may* or *may not* be from the same individual. That is a sticky point to which we will return in force in Chapter 3.

MNI seems to have originated in paleontology with individuals such as Stock (1929) and Howard (1930). It has been suggested that MNI was introduced to (zoo)archaeologists in 1953 by Theodore White (Grayson 1979), a paleontologist who worked with zooarchaeological collections, and this could well be correct. However, an archaeologist working a few years prior to White used MNI as a measure of taxonomic abundances.

In his unpublished Master's thesis, William Adams (1949:23–24) estimated the "approximate number of animals represented by the sample" of bones and teeth he

Table 2.4. *Some published definitions of MNI*

1. "the number of similar parts of the internal skeleton as for example the skull, right ramus of mandible, left tibia, right scaphoid" (Stock 1929:282).
2. "for each species, the left or the right of the [skeletal] element occurring in greatest abundance" (Howard 1930:81–82).
3. "the bone with the highest total will indicate the minimum number" (Adams 1949:24).
4. "separate the most abundant element of the species into right and left components and use the greater number as the unit of calculation" (White 1953a:397).
5. "the [skeletal] element present most frequently" (Shotwell 1955:330); "that number of individuals which are necessary to account for all of the skeletal elements (specimens) of a particular species found in a site" (Shotwell 1958:272).
6. the number of lefts and of rights of each element, those matching in terms of age and size tallied as from the same individual, those not matched tallied separately as from different individuals (Chaplin 1971).
7. "equal to the greatest number of identical bones per taxon" (Mollhagen et al. 1972:785).
8. "a count of the most frequent diagnostic skeletal part" (Perkins 1973:368).
9. "the most frequently occurring bone" (Uerpmann 1973:311).
10. "the number that is sufficient to account for all the bones assigned to the species; the most abundant body part" (Klein 1980:227).
11. "the least number of carcasses that could have produced the recovered remains . . . determined by taking the raw count of the most commonly retrieved bone element that occurs only once in the skeleton" (Gilbert and Singer 1982:31–32).
12. "may be based upon counts of the most abundant element present from one side of the body or on counts determined by joint consideration of skeletal parts represented; the size, age, and wear-state of specimens" (Badgley 1986:329).
13. "essentially the sample frequency of the most abundant skeletal part" (Plug and Plug 1990:54).
14. "the smallest number of individual animals needed to account for the specimens of a taxon found in a location" (Ringrose 1993:126).
15. "the most frequently occurring element" (Rackham 1994:39).
16. "the higher of the left- and right-side counts (if appropriate – obviously not if the most abundant element is an unpaired bone such as the atlas) is taken as the smallest number of individual animals which could account for the sample" (O'Connor 2000:59).

had studied. He chose "certain bones as readily identifiable, easily distinguished with regard to right or left position in the body and not commonly used for artifacts," and tallied up the occurrences of each for two taxa in each of five distinct recovery proveniences (p. 24). He noted that "since any one animal can possess only one of each of these bones, then the bone with the highest total will indicate the minimum number of mammals represented by the bone sample from that [recovery provenience]" (p. 24). Adams summed the MNI values indicated by each assemblage of bones from a unique recovery provenience and noted that in so doing, he had to assume "that parts of one individual are not represented from more than one [recovery provenience]" (p. 24); he assumed that skeletal remains in one provenience were independent of those in all others. Despite these significant insights, Adams (1949:24) abandoned MNI values because they provided *only* "minimum numbers," and he believed that assigning a "maximum number would be a matter of guesswork." Adams desired a quantitative unit that provided ratio–scale taxonomic abundances. Adams did not reference Stock or any other paleontologist who had previously used MNI as a quantitative unit. Circumstantial evidence therefore suggests that Adams invented (if you will) MNI independently of its invention in paleontology. But because he did not publish his discussion, few zooarchaeologists seem to have been aware of the MNI quantitative unit prior to White's work. At least few of them used MNI prior to the late 1950s, by which time it had been used in clever ways by Theodore White, who published his results in archaeological venues.

Unlike Adams, White (1953a, 1953b) did not seek to estimate taxonomic abundances when he introduced MNI to zooarchaeologists. Rather, he sought to estimate the amount of meat provided by each taxon; that was his target variable. He (1953a:397) noted that "four deer [*Odocoileus* sp.]" were needed to provide as much meat as "one bison [*Bison bison*] cow," and NISP would not reveal how much meat was provided by each of these taxa. Being a paleontologist who likely was familiar with, and used to seeing MNI values reported in the paleontological literature (e.g., Howard 1930; Stock 1929), White was aware of a quantitative unit (MNI) that could be easily converted to meat weight. White may have believed that there was no other reason that animal remains would be of interest to archaeologists, other than to reveal some aspects of human behavior. Diet – what folks ate – was an obvious human behavior reflected by animal remains.

Methods, including White's, to estimate meat weight (and the related variable, biomass) are discussed in Chapter 3. The important point here is that MNI was introduced to zooarchaeologists not as a replacement for NISP as a measure of taxonomic abundances. Rather, MNI was introduced to zooarchaeology in order to measure something else, specifically the amount of meat represented by a collection

of faunal remains. From a historical perspective, this is interesting for the simple reason that MNI was used in yet another discipline originally to estimate taxonomic abundances. The measurement of biomass and meat weight was introduced in that discipline as a replacement for MNI as a measure of diet, that is, for virtually the same reason MNI was introduced in zooarchaeology.

One of the things that ornithologists are interested in is the diet of birds. Raptors (hawks and eagles) and owls hunt, among other animals, small mammals – various insectivores, rodents, and leporids – and, depending on the taxon of the bird, swallow partial or complete carcasses of their prey. After 12–24 hours or so, a "pellet" of hair, bones, and teeth is regurgitated (the common terms in the ornithological literature are "egested" or "cast"). Depending on the bird, the bones and teeth are often in very good condition (not broken or excessively corroded from digestive acids) and retain many taxonomically diagnostic features (Andrews 1990). Because pellets are deposited beneath roosts (resting areas) and nests, a collection of pellets from such a location can reveal much about the diet of the bird.

Studies of such pellets and the faunal remains they contain have a deep history in ornithology (e.g., Errington 1930; Fisher 1896; Marti 1987; Pearson and Pearson 1947). Because ornithologists study the remains of prey in those pellets in order to answer some of the same questions that paleozoologists do (Which taxon is most abundant and which is least abundant on the landscape? Which taxon provided the most sustenance to the predator? [e.g., Andrews 1990; Mayhew 1977]), ornithologists have grappled with some of the same issues that paleozoologists have, especially with respect to how to quantify the remains of vertebrate prey. Ornithologists quickly figured out that NISP might not give a valid indication of which prey taxon was the most frequently consumed, so they did one of two things. They either tallied only skulls, or they determined the MNI based on whether the skull, left mandible, or right mandible was the most common skeletal element in a collection. They used both of these approaches as early as the 1940s (references in Lyman et al. 2003), describing how they counted taxonomic abundances. The earliest formal definition of MNI by an ornithologist of which I am aware is Mollhagen et al.'s (1972:785): the "minimum number of animals [is] equal to the greatest number of identical bones per taxon." No ornithologist who uses MNI, references Stock or any other paleontologist who used MNI, suggesting yet another independent invention of MNI. Near universal adoption of MNI as the quantitative unit of choice of ornithologists lead quickly to recognition that an MNI of five for each of two taxa did not give an accurate measure of diet when individuals of those two taxa were of rather different size. Thus, ornithologists determined the live weights of average adult individuals of common prey species and

used those data to determine the composition of a bird's diet (Steenhof 1983), much as White (1953a, 1953b) had done 30 years earlier.

Given that three separate disciplines have used MNI, and all of them (granting Adams's flirtation with it) seem to have independently invented it, one might think that MNI is a well-understood unit of measurement. It is commonsensical to calculate, and it has a basis in the empirically verifiable reality of the individuality and physical discreteness and boundedness of every organism. But MNI is not a well-understood quantitative unit. It has a number of problems, just like NISP. And also just like NISP, several of the problems with MNI are trivial or easily dealt with analytically, but one of them is rather serious.

Strengths(?) of MNI

Klein (1980:227) stated that unlike NISP, MNI is not affected by differential fragmentation, and suggested that this was a reason to seriously consider using MNI values as measures of taxonomic abundance, particularly when comparing assemblages with different degrees of fragmentation. He was concerned that a taxon, the remains of which had not been broken, would be underrepresented by NISP relative to a taxon the remains of which had been broken, all else being equal. Although Klein is correct that fragmentation will increase NISP, he is only partially correct because in reality fragmentation can influence MNI in two ways. First, fragmentation of moderate intensity, say, breaking each element into two more or less equal size pieces, will not influence MNI because specimens will retain anatomically and taxonomically diagnostic features (Lyman 1994b). Second, as the intensity of fragmentation increases, meaning that as fragments get smaller and represent less of the skeletal element from which they originate, the more difficult it will be to identify those fragments as to skeletal element represented and to taxon. This is so because progressively smaller fragments are successively less likely to retain anatomically and taxonomically diagnostic features (Lyman and O'Brien 1987). Thus fragmentation first increases NISP (but not MNI), but then as fragmentation intensifies, NISP decreases and so too does MNI.

The relationship between fragmentation and NISP, and that between fragmentation and MNI, was spelled out by Marshall and Pilgrim (1993) with respect to measuring the frequencies of individual skeletal parts. Because MNI is based on the most frequent skeletal part, Marshall and Pilgrim's findings are equally applicable to both NISP and MNI. Fragmentation influences MNI, although in a manner different than it does NISP. Breaking a skeletal element into pieces will first increase,

and then decrease NISP as fragmentation intensity increases and specimens become unidentifiable; moderate breakage will not influence MNI, but intensive fragmentation will result in a decrease in MNI as progressively more specimens fail to retain sufficient anatomical landmarks to allow identification.

Klein (1980) is correct that MNI will not be influenced by fragmentation, whereas NISP will be, but only in the limited case of assemblages with little fragmentation. Both NISP and MNI values will be influenced by *intensive* fragmentation. We do not know, however, the degree of fragmentation intensity at which fragmentation begins to decrease identifiability and to reduce values of both NISP and MNI (Lyman 1994b). The difficulty of establishing this degree of fragmentation is exacerbated by the fact that small fragments are unidentifiable to skeletal element.

MNI overcomes such problems as intertaxonomic variation in the number of identifiable elements per individual. But as I noted, in regard to this problem as it pertains to NISP, it is easy to analytically correct for intertaxonomic variation in the number of identifiable skeletal elements per individual. Either normalize all NISP by dividing those values by a value that accounts for variation in identifiable elements per skeleton, or delete from tallies those taxonomically unique elements (such as upper incisors in equids when comparing their abundance to bovids, who don't have upper incisors).

The most important advantage of MNI is that it overcomes possible specimen interdependence because of how MNI is defined (Table 2.4). As Ringrose (1993:127) noted, the "basic principle of the MNI is to avoid 'counting the same animal twice.'" If MNI is derived according to the definitions in Table 2.4, there is no way for two or more specimens in one assemblage to be from the same individual. This is illustrated in Table 2.5. Notice in that table of fictional data that there is a minimum of seven individuals (= MNI). Notice also that all postcranial elements have been assigned to an individual already represented by a skull and both mandibles. The left scapulae do not represent, so far as we can tell in light of modern methods, an eighth individual, a ninth individual, and so on, nor do the right scapulae, the left humeri, and so on. Here, $\sum$NISP = 71, but obviously if we tally NISP, we count the first individual fifteen times (= $\sum$NISP for that individual), the second individual is counted fourteen times, and so on. Of course, the right radius assigned to individual number five may actually go with individual number four, or with individual number eight, but we cannot determine that. Thus, the number of individuals represented by the seventy-one specimens listed in Table 2.5 is a minimum of seven individuals, but the specimens might represent more.

MNI would seem to avoid the interdependence problem *within an assemblage* such as shown in Table 2.5. But note the emphasized phrase – *within an assemblage.*

Table 2.5. *A fictional sample of seventy-one skeletal elements representing a minimum of seven individuals (I)*

	I.1	I.2	I.3	I.4	I.5	I.6	I.7
Skull	+*	+	+	+	+	+	+
L mandible	+	+	+	+	+	+	+
R mandible	+	+	+	+	+	+	+
L scapula	+	+	+	+	+	+	
R scapula	+	+	+	+	+	+	
L humerus	+	+	+	+	+	+	
R humerus	+	+	+	+	+		
L radius	+	+	+	+	+		
R radius	+	+	+	+	+		
L innominate	+	+	+	+			
R innominate	+	+	+	+			
L femur	+	+	+				
R femur	+	+	+				
L tibia	+	+					
R tibia	+						
∑NISP	15	14	13	11	9	6	3

* + denotes a skeletal element is represented.

What about *between* assemblages? What if the left femur of an individual is in one assemblage and the matching right femur is in another assemblage? Were we to tally each as part of an MNI calculation in the two assemblages, we would have counted that single individual twice. This introduces the most serious problem with MNI – aggregation – and there are other, less serious but significant problems as well.

Problems with MNI

As with NISP, analysts have identified what they take to be serious problems with MNI (e.g., Casteel n.d.; Fieller and Turner 1982; Gilbert et al. 1981; Grayson 1973, 1979, 1984; Klein 1980; Klein and Cruz-Uribe 1984; Plug and Plug 1990; Ringrose 1993; Turner 1983, 1984; Turner and Fieller 1985). These include:

1 MNI is difficult to calculate because it is not simply additive;
2 MNI can be derived using different methods, thereby reducing comparability;
3 MNI values do not accurately reflect the thanatocoenose or the biocoenose;

4 MNI values exaggerate the importance of rarely represented taxa, or taxa represented by low NISP values;
5 MNI values are minimums and thus ratios of taxonomic abundances cannot be calculated;
6 MNI is a function of sample size or NISP, such that as NISP increases, so too does MNI; and
7 different aggregates of specimens comprising a total collection will produce different MNI values.

The first problem – MNI is difficult to determine – is not worth considering. No one has ever said research of any kind was, or should be, easy. But related to this problem is the fact that MNI is not additive like NISP is. Rather, every time a new bag of faunal remains is opened and specimens identified, one has to rederive the MNI (assuming it was derived before). This is so because the most common skeletal part per taxon may change with the addition of another bag of bones. This problem, too, is trivial given that research often involves calculation and recalculation, again and again, as new data are collected or as new insights are gained and adjustments are made to a data set.

That different researchers use different methods to derive MNI values (the second problem) is evidenced by variations in the definitions of MNI in Table 2.4. This problem is akin to the one that different analysts will produce different NISP values for the same collection of remains given their varied expertise at identification. There is no real way to control this problem, so the methods used to derive MNI values should be stated explicitly. Were remains matched by size? By age? By recovery context? All of these or other variables? Potential and varied use of these criteria make MNI a *derived* measure.

Some have argued that to not attempt to match potentially paired remains such as left femora with right femora – to determine if they originated in the same individual – results in a misleading MNI value (Fieller and Turner 1982). White (1953a:397) thought that matching would require "the expenditure of a great deal of effort with small return [to] be sure all of the lefts match all of the rights." That is, he believed that checking every possible bilateral pair of bones for matches (based on the notion of bilateral symmetry – that a left element is a mirror image of its right element mate), and assuming that each of those matches derived from the same individual – would result in a relatively small increase in the MNI value, and thus not much in the way of alteration of taxonomic abundances. Whether or not White was correct with respect to the magnitude of change in MNI values when matching is undertaken is unclear, but likely is assemblage specific for many reasons (Lyman 2006a). Some argue that

the difference would be considerable whereas others seem to think it would not be if "a great deal of effort" were expended, whatever the result, it will be a function of the identified assemblage at hand and the procedures used to analytically manipulate the bilateral pair data rather than the act of matching and identifying bilateral pairs itself (compare Fieller and Turner [1982] with Horton [1984]).

The third problem, too, can be said to characterize both NISP and MNI. NISP is a count of identifiable specimens in the assemblage rather than a measure of the thanatocoenose or the biocoenose. Similarly, given its definition, MNI is the *minimum* number of animals necessary to produce the identified specimens comprising the identified assemblage. Both quantitative units *describe* the assemblage using two rather different variables. Whether either NISP or MNI more accurately reflects the thanatocoenose or the biocoenose cannot be assumed given the contingent and particularistic nature of the taphonomic history of the assemblage (e.g., Gilbert and Singer 1982; Ringrose 1993; Turner 1983). But, as I noted with respect to NISP, there are ways to test if the identified assemblage rendered as a set of MNI values accurately reflects the biocoenose. Are the taxonomic abundances indicated by MNI values what would be expected given independent evidence of environmental conditions? Do MNI abundances match those from contemporaneous nearby faunal assemblages that experienced independent taphonomic histories? If the answers to these questions are all "yes," then it would be reasonable to suppose that the taxonomic abundances in the collection under study are fairly accurate reflections of those abundances in the thanatocoenose as well as the biocoenose.

The fourth problem can be appreciated by recalling Table 2.3. There, taxon A is rarely represented (NISP = 1) whereas taxon B is frequently represented (NISP = 10), but both taxa have an MNI of 1. MNI exaggerates the representation of taxon A relative to taxon B's representation. Although this observation is true, it is merely the converse of the related problem with NISP. Thus one might argue that MNI and NISP are equally flawed in this respect. If you are uncomfortable with that, you could note that taxon A in Table 2.3 is represented by one tibia, and tally only tibiae identified as taxon B when comparing abundances of these two taxa. It is likely that such a procedure would decrease the disparity in representation of the two taxa, but it also demands the assumption that the other specimens of taxon B are interdependent with that taxon's tibiae, an assumption that would likely be difficult to warrant empirically or theoretically. (Some specimens may be interdependent, but it seems improbable that all would be interdependent.)

Plug and Plug (1990) identified the fifth problem: MNI values are minimums and thus ratios of MNI values cannot be validly calculated (see also Gilbert et al. 1981). They note that if the MNI of taxon *A* is 10 and the MNI of taxon *B* is 20, we cannot

use simple arithmetic to calculate the *A*:*B* ratio because, with respect to the true number of individuals, it is very likely that $A \geq 10$ and $B \geq 20$. Thus any ratio *A*:*B* cannot be validly calculated. MNI values are not ratio scale. Instead, they are *perhaps* ordinal scale. Thus we can say, in Plug and Plug's case of *A*:*B*, that it is likely that $A < B$, but we cannot say by how much given that both *A* and *B* are minimum values, and we don't know their true values. A similar argument can be made with respect to NISP values. They are maximum estimates of taxonomic abundances, so a ratio of NISP values for two taxa, although easily calculated, may not actually be a ratio scale measure of taxonomic abundances. That MNI values are minimums is clear from how they are derived (Table 2.4). Paleozoologists have known these things since the MNI quantitative unit was introduced (e.g., Adams 1949). What is perhaps less well-known is that recent simulations indicate that MNI often provides values considerably lower than the actual number of individuals (ANI) present in a collection (Rogers 2000a). To illustrate this, look at Table 2.5 one more time. Here, the MNI is seven; the NISP is seventy-one. If each skeletal element is independent of every other element (each comes from a different organism), then the ANI is seventy-one, an order of magnitude greater than the MNI value.

The fact that MNI increases as NISP increases (problem 6) has been recognized for some time in zooarchaeology (Casteel 1977, n.d.; Ducos 1968; Grayson 1978a). Some have argued that this statistical relationship warrants use of NISP rather than MNI to measure taxonomic abundances; the reason is that the same information regarding taxonomic abundances is contained in both quantitative units, so there is no reason to determine MNI. Others have noted that although this statistical relationship does indeed exist between the two measures, the precise nature of the relationship depends on the particular set of remains involved (e.g., Bobrowsky 1982; Grayson 1984; Hesse 1982; Klein and Cruz-Uribe 1984). Some researchers use the last observation – that the relationship between NISP and MNI is statistically particularistic – to argue that one cannot predict MNI from NISP in a new sample based on the statistical relationship of the two in previously studied samples, so perhaps MNI should be determined (Klein and Cruz-Uribe 1984). This is a clever insight, and it is correct, but it does not mean we must determine MNI values when seeking measures of taxonomic abundance. We need not determine MNI because of the relationship between the NISP for a taxon in an assemblage and its attendant MNI.

It is commonsensical that as the NISP of a taxon increases so too should the MNI for that taxon. This is so because every individual skeleton comprises a limited number of elements (or what might become identifiable specimens comprising the paleozoological record). Adding randomly chosen skeletal elements selected from, say, 100 skeletons of an identified assemblage, the first element will contribute one individual.

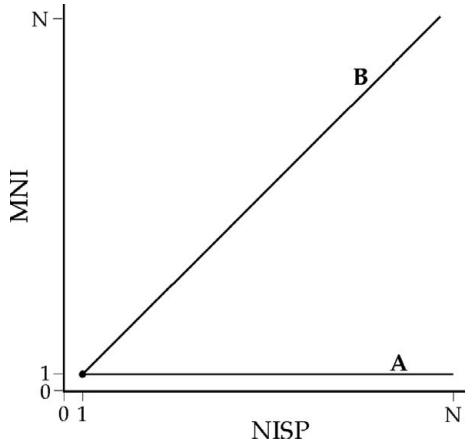


FIGURE 2.3. The theoretical limits of the relationship between NISP and MNI. Modified from Grayson (1978a). Line A indicates that every new specimen does not contribute a new individual. Line B indicates that every new specimen contributes a new individual.

That is, NISP = 1 = MNI. The second element will contribute another NISP ($\sum$NISP = 2), but that element might, or might not, contribute another individual ($\sum$MNI = 1 or 2). The third element will contribute yet another NISP ($\sum$NISP = 3), and it might contribute another individual or it might not ($\sum$MNI = 1, 2, or 3). And so on until the probability of adding a bone or tooth of an already represented skeleton is greater than the probability of adding a bone of an unrepresented skeleton, at which point the rate of increase in MNI will slow relative to the rate of increase of NISP. As Grayson (1978a) noted, there are two limits to the possible relationship between NISP and MNI. Either every new skeletal element derives from the same individual and thus every NISP > 1 contributes nothing to the MNI tally, or every new skeletal element derives from a different, unique individual and thus every NISP $\geq$ 1 contributes another MNI. These relationships express the limits of all possible relationships between NISP and MNI (Figure 2.3).

The individual limits to the relationship between NISP and MNI (Figure 2.3) are unlikely to be found in the real world. Unless one is dealing with, say, the moderately fragmented skeleton of a single individual animal (NISP is several hundred), it is likely that NISP will increase more rapidly than MNI. In fact, unless one is dealing with an assemblage of remains of a single taxon that has but one identifiable skeletal element (such as the unbroken shells of a gastropod), it is likely that new NISP will often be added without adding any MNI. The general relationship between NISP and MNI is described by the line in Figure 2.4. It is relatively easy to show that this is indeed the relationship that is found in case after case. The relationship

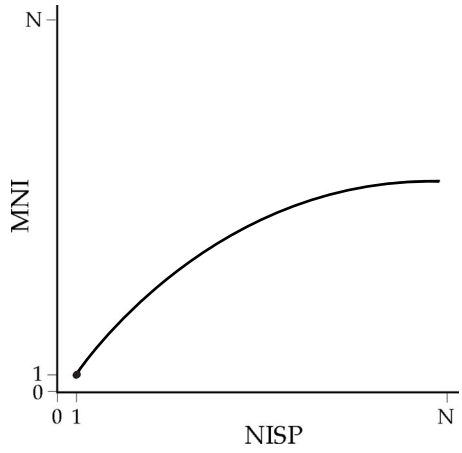


FIGURE 2.4. The theoretically expected relationship between NISP and MNI. Modified from Casteel (1977). As NISP increases, it takes progressively more specimens to add new individuals.

is curvilinear because, given a finite NISP and a finite MNI, specimens from an individual already represented are progressively more likely to be added as the sample size ($\sum$NISP) increases. Whatever the kind of skeletal part that is the most frequent or most common and thus defines MNI, that kind of part will become progressively more difficult to find (Grayson 1984).

Ducos (1968) found that this curvilinear relationship could be made linear (and thus perhaps more easily understood when graphed) if both the NISP data and the MNI data were log transformed (see Box for additional discussion). As Grayson (1984:52) later noted, untransformed NISP data and untransformed MNI data may sometimes be related in a linear fashion, but they often are not. The means to tell if they are not involves examination (either visual or statistical) of the residuals (the distance above and below the regression line of the plotted points and the pattern of the distribution of those points). Typically, log transformation reduces the dispersal of points to a statistically insignificant level. The slope of the best-fit regression line summarizes the rate of change in MNI relative to the rate of change in NISP and is described by a single number representing a power function (or exponent); the larger the number, the steeper the slope.

Casteel (1977) found the relationship modeled in Figure 2.4 in a series of assemblages of zooarchaeological and paleontological materials representing numerous taxa. His data originally were comprised of 610 paired NISP–MNI values. (A *paired NISP–MNI value* is the NISP value and the MNI value for a taxon in an assemblage of remains.) Casteel subsequently expanded his data set (Casteel n.d.) to include 3,440

BOX 2.1

It is often easier to grasp intuitively a linear relationship between two variables than a curvilinear one. In many cases log transformation of NISP and MNI data causes what is otherwise a curvilinear relationship to become linear. The typical form of a linear relationship can be expressed by the equation $Y = a + bX$, where X is the independent variable (in this case, NISP), Y is the dependent variable (in this case, MNI), a is the Y intercept (where the line describing the linear relationship intersects the Y axis), and b is the slope of the line (where the slope of the line describing the linear relationship represents how fast Y changes relative to change in X). The simple best-fit regression line in a graph showing the relationship between log transformed NISP data and log transformed MNI data is described by the formula $Y = aX^b$, where the variables Y, a, X, and b are as defined above. This formula describes what is referred to as a power curve; if b is positive the curve extends upward from the lower left to the upper right of the graph; if b is negative the curve extends downward from the upper left to the lower right. If we transform both sides of $Y = aX^b$ to logarithms, then we have $\log Y = \log a + b \log X$, linear relationship between $\log Y$ and $\log X$. In this volume, I present the relationship between $\log X$ or log NISP, and $\log Y$ or log MNI, in the form $Y = aX^b$, or what is simply a different form of the linear relationship. The Y intercept should be zero, given that a zero value for NISP must produce a zero value for MNI, but practice has been to allow the empirical data to identify a Y intercept; I follow this practice here noting that should the empirically determined Y intercept differ considerably from zero, the data used should be inspected to determine why. Variables a and b are constants determined empirically for each data set.

paired NISP–MNI values. (The manuscript in which Casteel used this larger data set was never published. It was written in 1977, and afterwards cited occasionally by his colleagues [e.g., Bobrowsky 1982; Grayson 1979]. I obtained a copy of the manuscript from Grayson in the late 1970s.) In both cases, Casteel found a statistically significant relationship between NISP and MNI like that shown in Figure 2.4. Bobrowsky (1982) found the same relationship between paired NISP–MNI values using much smaller data sets. Both Casteel (1977, n.d.) and Bobrowsky (1982) graphed the relationship using untransformed data; Grayson (1984) and Hesse (1982) used log-transformed data in their graphs of the relationship. Grayson (1984) summarized many cases that had been reported by others, and reported several new cases to show that the relationship was essentially ubiquitous. Klein and Cruz-Uribe (1984) found the same

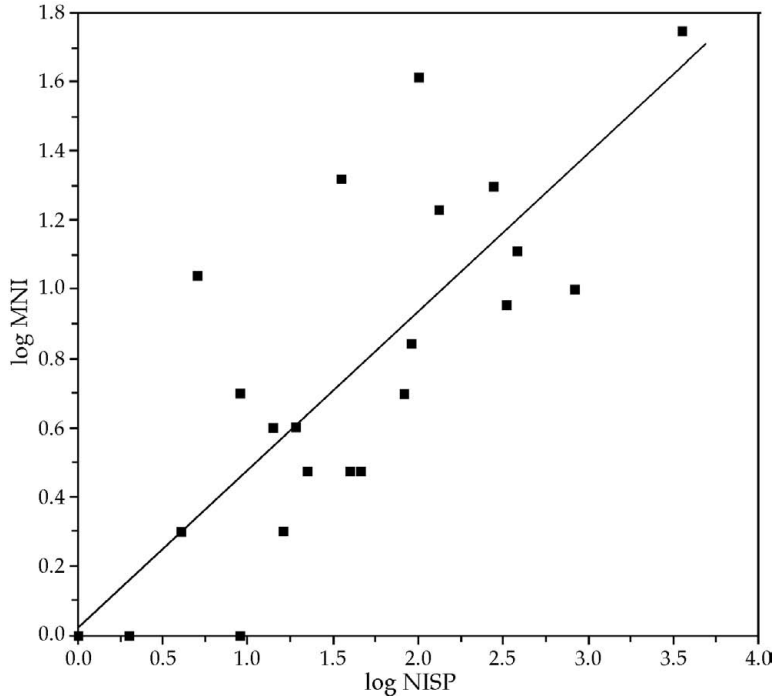


FIGURE 2.5. Relationship between NISP and MNI data pairs for mammal remains from the Meier site (data in Table 1.3). See Table 2.6 for statistical summaries.

relationship between NISP and MNI using data sets different from those used by Casteel, Bobrowsky, Grayson, and Hesse.

The large collection from Meier ($\sum$NISP = 5939) shows the relationship between NISP and MNI nicely (Figure 2.5). The NISP–MNI data pairs are strongly correlated (Pearson's $r = 0.8734$, $p < 0.0001$). The slope of the best-fit regression line (= 0.487) is similar to that reported by others for other data sets; Casteel (1977) reported a slope of 0.52, for example, and Grayson (1984) reported six others that ranged from 0.40 to 0.64. The precontact assemblage from Cathlapotle (Table 1.3) also shows the nature of the relationship between NISP–MNI data pairs (Figure 2.6), as does the postcontact assemblage from that site (Figure 2.7). In all three cases, the correlation coefficient is strong ($r > 0.87$) and significant ($p < 0.0001$). The statistical relationships between the two variables in each of these three assemblages are summarized in Table 2.6.

The relationship between NISP and MNI shown in Figures 2.5, 2.6, and 2.7 is not unique to the Portland Basin. Recall that Casteel, Grayson, Hesse, Bobrowsky, and Klein and Cruz-Uribe found exactly the same relationship between the two variables in collections from all over the world and representing many time periods and taxa.

Table 2.6. *Statistical summary of the relationship between NISP and MNI for mammal assemblages from Meier (Figure 2.5) and Cathlapotle (Figures 2.6 and 2.7) (see Table 1.3 for data).*

Site	Regression equation	*r*	*p*	∑NISP	N of taxa
Meier	$MNI = -0.06(NISP)^{0.487}$	0.873	<0.0001	5,939	26
Cathlapotle, precontact	$MNI = -0.098(NISP)^{0.42}$	0.916	<0.0001	2,372	21
Cathlapotle, postcontact	$MNI = -0.0557(NISP)^{0.44}$	0.901	<0.0001	3,834	24

Together, these cases suggest that the relationship is nearly ubiquitous. Table 2.7 summarizes the statistical relationship between NISP and MNI in fourteen assemblages of mammal remains from fourteen sites in eastern Washington State. I, along with two fellow graduate students at the time, identified the taxa in these collections in the late 1970s. All fourteen collections display the same kind of relationship between NISP and MNI as is evident for Meier and Cathlapotle.

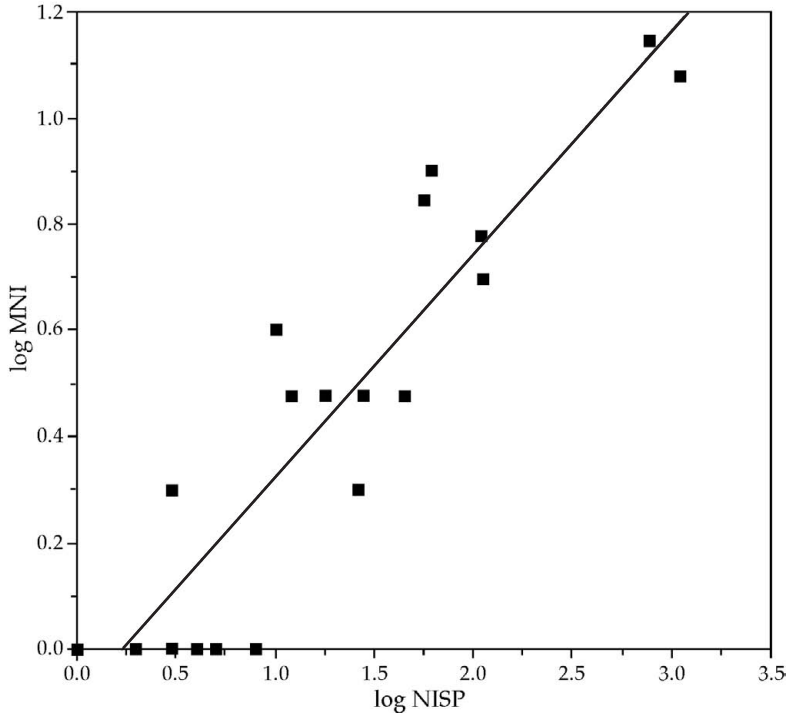


FIGURE 2.6. Relationship between NISP and MNI data pairs for the precontact mammal remains from the Cathlapotle site (data in Table 1.3). See Table 2.6 for statistical summaries.

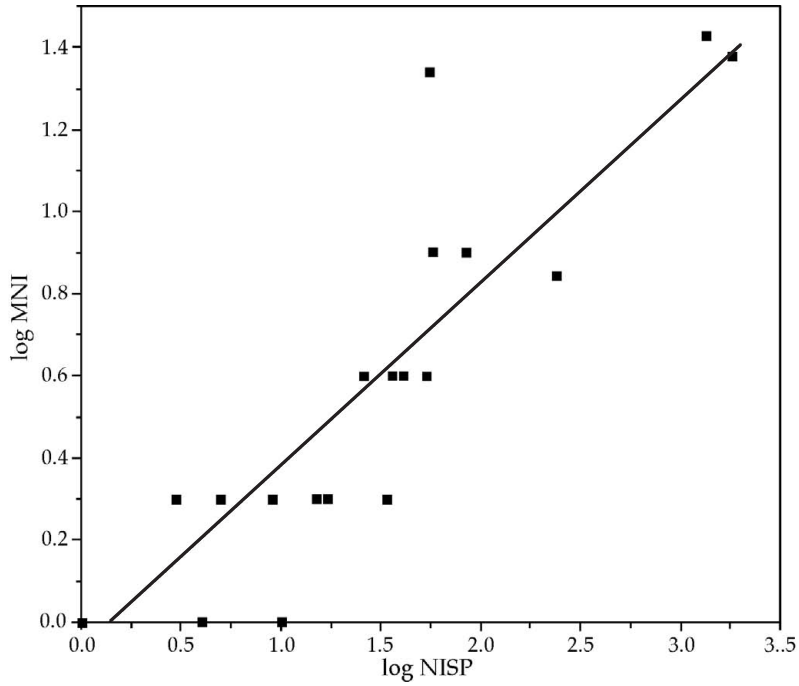


FIGURE 2.7. Relationship between NISP and MNI data pairs for the postcontact mammal remains from the Cathlapotle site (data in Table 1.3). See Table 2.6 for statistical summaries.

That NISP–MNI data pairs are often tightly related (the correlation coefficient is large) even in nonarchaeological contexts is also easy to show. Consider a sample of eighty-four pellets likely cast by a barn owl (*Tyto alba*) in eastern Washington State. NISP and MNI values for the total assemblage of prey remains in the eighty-four pellets (Table 2.8) are arrayed in a bivariate scatterplot in Figure 2.8. The relationship between the values of the two is linear and strong ($r = 0.989$, $p = 0.0002$), as it is among the archaeological samples discussed previously; the regression equation is: $\text{MNI} = -1.57(\text{NISP})^{0.935}$. The same relationship holds for paleontological collections as well, as Grayson (1978a) showed some years ago.

The fact that NISP and MNI are often strongly correlated could be used to argue that we should use MNI as the quantitative unit for measures of taxonomic abundance because of the potential for the interdependence of specimens in NISP tallies. Indeed, Hudson (1990) argued on the basis of ethnoarchaeological data, and Breitburg (1991) on the basis of historic-era zooarchaeological data supplemented by written documents, that MNI provides more accurate measures of taxonomic abundances than NISP. That MNI would indeed sometimes be a more accurate measure of taxonomic abundances than NISP is to be expected given everything we know about the two quantitative units and the influences of taphonomic processes and recovery

Table 2.7. *Statistical summary of the relationship between NISP and MNI for mammal assemblages from fourteen archaeological sites in eastern Washington State*

Site	Regression equation	r	p	∑NISP	N of taxa
45DO273	$MNI = -0.114(NISP)^{0.58}$	0.816	0.0136	84	8
45OK2A	$MNI = 0.01(NISP)^{0.36}$	0.849	0.0019	366	10
45DO282	$MNI = -0.178(NISP)^{0.628}$	0.875	0.0004	426	11
45DO211	$MNI = -0.12(NISP)^{0.64}$	0.847	< 0.0001	474	15
45DO285	$MNI = -0.19(NISP)^{0.51}$	0.721	0.0024	491	15
45DO214	$MNI = -0.07(NISP)^{0.44}$	0.765	0.0003	536	17
45DO326	$MNI = 0.02(NISP)^{0.3}$	0.56	0.0242	640	16
45DO242	$MNI = -0.093(NISP)^{0.4}$	0.89	< 0.0001	673	13
45OK287	$MNI = -0.04(NISP)^{0.21}$	0.786	0.007	807	10
45OK250	$MNI = -0.019(NISP)^{0.41}$	0.776	0.003	1,077	12
45OK4	$MNI = -0.072(NISP)^{0.48}$	0.881	< 0.0001	1,108	15
45OK2	$MNI = -0.158(NISP)^{0.4}$	0.769	0.0002	2,574	18
45OK11	$MNI = -0.124(NISP)^{0.5}$	0.849	< 0.0001	3,549	24
45OK258	$MNI = -0.094(NISP)^{0.47}$	0.863	< 0.0001	4,433	21

and identification skills. What we don't know, and can't really know most of the time, is whether MNI is a more accurate measure of taxonomic abundances than NISP for any given assemblage of paleozoological remains. Hudson (1990) and Breitburg (1991) knew that MNI provided more accurate measures of the thanatocoenosis because they knew the original taxonomic abundances of the death assemblage. We don't

Table 2.8. *Maximum distinction (each pellet considered independently) and minimum distinction (all pellets considered together) MNI values for six genera of mammals in a sample of eighty-four owl pellets. The right femur is the most abundant element for* Microtus*, and the left femur is the most abundant element for* Peromyscus *in the minimum distinction column*

Taxon	NISP	Minimum distinction MNI	Maximum distinction MNI
Sylvilagus	5	1	2
Reithrodontomys	19	5	5
Sorex	46	5	7
Thomomys	68	8	12
Microtus	705	104	118
Peromyscus	1,266	188	220
∑	2,109	310	364

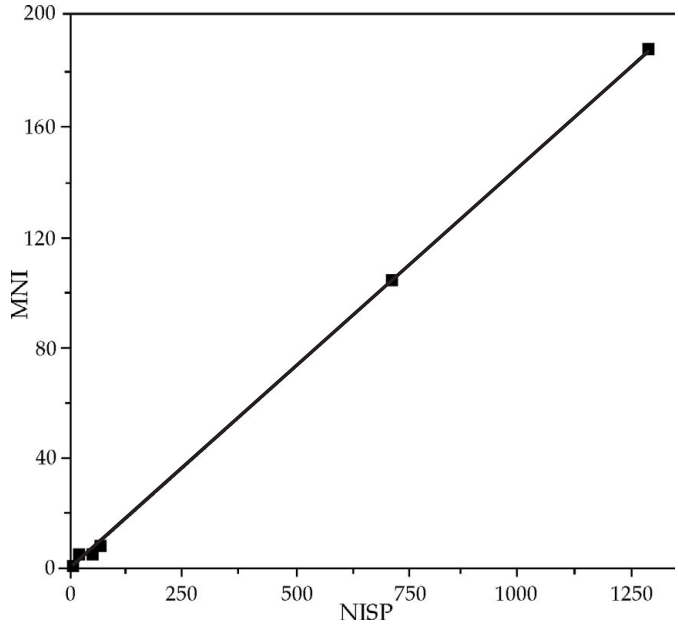


FIGURE 2.8. Relationship between NISP and MNI data pairs for remains of six mammalian genera in eighty-four owl pellets. The regression equation is $MNI = -0.389(NISP)^{0.149}$, and the correlation coefficient of the simple best-fit regression line is significant ($r = 0.999$, $p < 0.0001$). Data from Table 2.8.

know what the thanatocoenosis is in paleozoological contexts; it might be our target variable.

That the relationship between NISP and MNI is particularistic and its precise nature is dependent on the samples used is true. But the truth of that claim is not a necessary basis for rejecting NISP in favor of MNI. This is so for several reasons. First, as I noted earlier, NISP often contains virtually the same information regarding taxonomic abundances as does MNI. Second, there are fewer analytical steps in tallying NISP than in deriving MNI, so there are fewer layers (to borrow a metaphor) in the house of cards upon which NISP rests than in the house of cards upon which MNI rests. Do not misinterpret this second point; the simplest method is not being advocated as the best just because it is simpler or easier or contains fewer analytical steps. Rather, because NISP contains fewer steps and, more importantly, fewer assumptions than MNI regarding taphonomy, recovery, identification skills, and the like, perhaps NISP should be preferred. Finally, there is still problem seven, which concerns how to aggregate faunal remains in order to produce MNI values. No such problem exists with NISP, reflecting the fact that NISP is simply additive and that MNI is not additive. It is time, then, to turn to what is likely the most serious problem with MNI.

Table 2.9. *Adams's (1949) data for calculating MNI values based on* Odocoileus *sp. remains. MNI-I is equivalent to the MNI minimum distinction (one aggregate, MNI = 118); MNI-II is equivalent to the MNI maximum distinction (five aggregates, MNI = 120)*

Recovery provenience	Distal right humerus	Distal left humerus	Distal right femur	Distal left femur	MNI-II
A	38	42	10	11	42
B	5	12	1	1	12
C	30	42	10	10	42
D	11	9	3	5	11
E	9	13	2	7	13
MNI-I	93	118	26	34	

Aggregation

Although MNI solves the problem of interdependence of specimens inherent to NISP, MNI has its own significant problem. That problem is readily introduced by considering the data that Adams (1949) presented when he determined the MNI of deer (*Odocoileus virginianus*) per recovery provenience unit in the collection he studied (Table 2.9). Adams distinguished what he referred to as "Minimums I" and "Minimums II," the individual column totals and the individual row totals, respectively. I have substituted MNI for "Minimums" in Table 2.9 because MNI is indeed what Adams meant. Notice that were one to ignore recovery provenience, and just tally up the most frequently occurring skeletal part, distal left humeri would be most abundant among the four skeletal parts, so the site-wide MNI – or Adams's "MNI-I" – is 118. But if the analyst were to tally up the most frequently occurring skeletal part per unique recovery provenience, then distal left humeri would be the most abundant skeletal part in four of the five recovery proveniences, but right distal humeri would be the most abundant skeletal part in the fifth recovery provenience. Thus the total MNI for the values summed over the five recovery proveniences – Adams's "MNI-II" – is 120. Grayson (1984) later presented an example from a single site in which differences between MNImni and MNImax values varied across nearly two dozen taxa from 0 to 250 percent (the latter, MNImin = 15 and MNImax = 38).

It makes little difference whether Adams's five distinct recovery proveniences are horizontally distinct (like units in an excavation grid), vertically distinct (as with strata), or both (as with grid units per stratum). His data illustrate the most significant problem that attends MNI. This problem was revealed by Adams (1949:24) when he commented that one must assume "that parts of one individual are not

represented from more than one [recovery provenience]." He said this with specific reference to his "MNI-II" values. But he only revealed the problem; he did not explore its implications. This problem and its implications were later documented at length by Grayson (1973, 1979, 1984). This problem is, in short, known as the aggregation problem, where an aggregate is an assemblage or collection of faunal remains the boundaries of which are chosen by the analyst, whether those boundaries correspond to stratigraphic boundaries or arbitrarily and artificially bounded excavation/collection units.

Grayson (1973) termed what Adams called "Minimums I" values the *minimum distinction method*, and termed what Adams called "Minimums II" values the *maximum distinction method*. The former involves determination of MNI for the complete assemblage considered as one aggregate; the latter involves determination of MNI independently for each assemblage, each from a distinct recovery provenience specified by the analyst. The minimum distinction method is so-called because it produces the lowest or smallest MNI values for a complete collection. The maximum distinction method is so-called because it produces the greatest or largest MNI values for a collection (MNI values for all assemblages from unique recovery proveniences are summed); it produces more than the minimum distinction method because it considers a large number of (small) aggregates (or [sub]assemblages of remains). The minimum distinction method considers only one large aggregate – all remains treated as a single collection. Adams did not care for either the minimum distinction method or the maximum distinction method because, despite the differences in their results, both produced *minimum* numbers of individuals. Furthermore, the maximum distinction method – determining MNI based on individual recovery proveniences – required that one assume specimens in one provenience unit were independent of all specimens in other provenience units, and Adams did not want to make that assumption. It is fitting that we hereafter refer to this potential problem of interaggregate interdependence of skeletal specimens as *Adams's dilemma*.

That the aggregation problem is widespread is easy to show. Recall the collection of remains of six genera of prey in eighty-four owl pellets; the NISP–MNI data pairs for this collection are plotted in Figure 2.8. That figure is based on the minimum distinction method because all remains were lumped together to form one aggregate. This means that only one skeletal element per taxon contributes to the MNI, regardless of how many pellets contain remains of a taxon. What happens to the MNI values for the taxa represented in the sample of eighty-four pellets when one shifts from the minimum to the maximum distinction method is shown in Table 2.8. The NISP values stay the same regardless of how MNI is determined – whether the maximum or minimum distinction method is used. The MNI is greater in five of six taxa

Table 2.10. *Differences in site total MNI between the MNI minimum distinction results and the MNI maximum distinction results*

Site	N of assemblages	NISP	Richness (N of genera)	MNImin	MNImax	N taxa increase	Mean increase per genus
45OK2A	4	366	10	30	39	4 of 10	0.9
45DO211	4	474	15	108	117	4 of 15	0.6
45DO285	4	491	15	66	102	12 of 15	2.4
45DO214	4	536	17	67	108	11 of 17	2.4
45DO326	4	640	16	53	81	14 of 16	1.75
45DO242	4	673	13	38	52	7 of 13	1.1
45OK250	3	1,077	12	62	79	8 of 12	1.4
45OK4	3	1,108	15	65	82	7 of 15	1.1
45OK2	4	2,574	18	66	105	13 of 18	2.2
45OK11	2	3,549	24	202	231	14 of 24	0.6
45OK258	2	4,433	21	117	139	10 of 21	1.0

when the maximum distinction method is used relative to the minimum distinction method (Table 2.8). The ratio of *Peromyscus* to *Microtus* – the subject of published interpretations of this collection (Lyman et al. 2001, 2003) – shifts from 1.81:1 for the minimum distinction MNI, to 1.86:1 for the maximum distinction MNI, to 1.80 for NISP. In this case, the differences are small, and statistically insignificant; the chi-square value is 0.41 if *Sylvilagus* is omitted so that the assemblage's data pairs meet the requirements of the test ($p > 0.5$). Even given the small differences in ratios of *Peromyscus* to *Microtus*, the critical question is: Which ratio is correct? There is no clear or obvious answer.

The aggregation problem is pernicious. Of the fourteen archaeological assemblages summarized in Table 2.7, eleven have multiple components or temporally distinct (sub)assemblages; the other three consist of only one assemblage. For purposes of generating the regression equations in Table 2.7, I used the minimum distinction MNI values for all fourteen sites. What happens to the MNI values for the eleven sites with multiple (sub)assemblages if the maximum distinction method is used and MNI is derived for each taxon in each (sub)assemblage independently? First, the total MNI for each of the eleven sites increases when one shifts from MNImin(imum distinction) to MNImax(imum distinction) (Table 2.10). Why? Because more kinds of most common elements are specified in the latter than in the former.

Second, the total MNI for each site increases between nine and forty-one when one shifts from MNImin to MNImax; the average increase is 23.7 individuals per

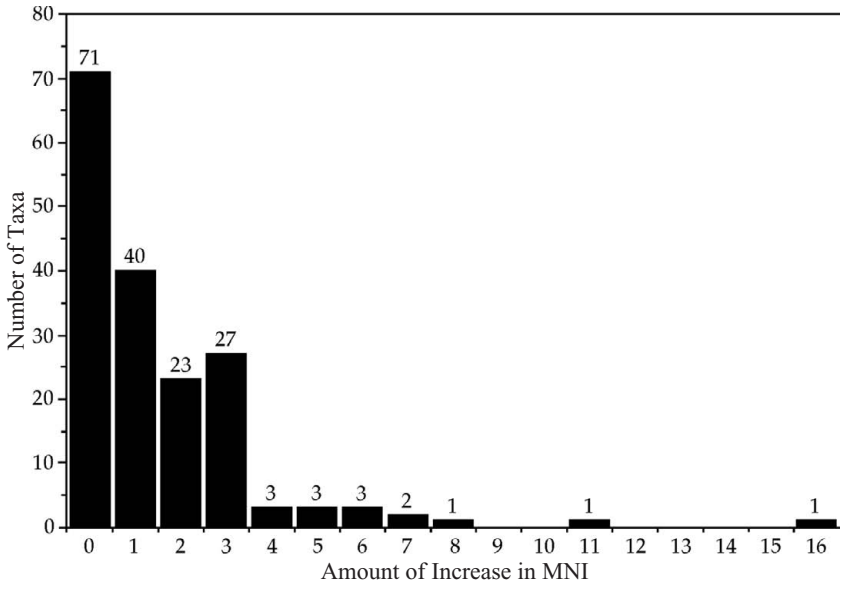


FIGURE 2.9. Amount by which a taxon's MNI increases if the minimum distinction MNI is changed to the maximum distinction MNI in eleven assemblages (see Table 2.10 for other data on these assemblages).

site (Table 2.10). This is not a lot in terms of absolute abundance, but think of it this way: In site 45OK2A, MNImin is thirty and MNImax is thirty-nine; that is a 30 percent increase. In the sample of eleven sites, the site total MNImax increases over MNImin from 8 percent (45DO211) to as much as 61 percent (45DO214); the average increase is a bit more than 35 percent. The third thing to note regarding the shift from MNImin to MNImax is that four to fourteen taxa per site increase in abundance. Not all taxa increase, and any given taxon does not increase consistently in all sites. Ratios of taxonomic abundances shift around rather unpredictably as a result. Note, for example, that the amount by which any taxon's MNI increases is one to sixteen (Figure 2.9). Consider, for example, how the ratio of deer (*Odocoileus* spp.) to gopher (*Thomomys* sp.) changes across all eleven sites when one uses MNImin compared to when one uses MNImax (Figure 2.10). If the MNI of both taxa changed consistently (say, all increase by 10 percent) when shifting from the minimum to the maximum distinction method, the ratios would not change and all points would fall on the diagonal in Figure 2.10. Instead, the eleven collections fall various distances from that line, meaning that the ratios change more in some sites than in others; the changes in most abundant skeletal parts are not patterned. There is marked variation in which skeletal part defines the MNI for either or both deer and gophers across the (sub)assemblages.

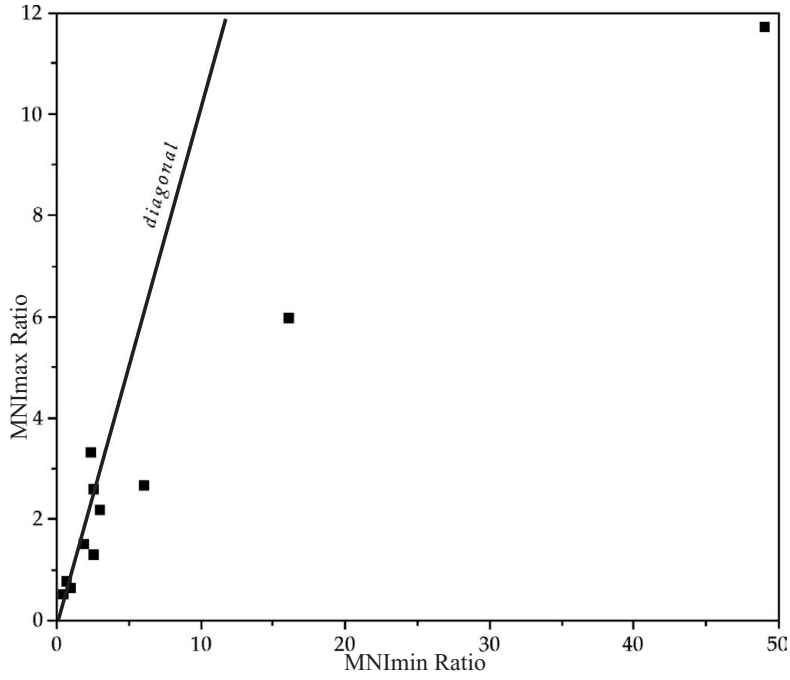


FIGURE 2.10. Change in the ratio of deer (*Odocoileus* spp.) to gopher (*Thomomys* sp.) abundances in eleven assemblages when MNImax is used instead of MNImin. If the ratios did not change, all points should fall on the diagonal line rather than above and to the left, or below and to the right of that line.

The most abundant skeletal part for each of the thirteen mammalian genera at Cathlapotle (Table 1.3) that have more than 1 MNI in each of the two temporally distinct (sub)assemblages are listed in Table 2.11. Only three taxa (of a possible thirteen) have more than one most abundant part (e.g., *Castor* in the postcontact assemblage), indicating rather skeletally uneven representation of individual carcasses. More importantly, it is virtually impossible to predict which part of a genus will be most abundant in one (sub)assemblage based on which part of that genus is most abundant in the other (sub)assemblage, particularly when left and right-side designations are considered (Table 2.11). If the side designation is not considered, then only four genera (*Aplodontia, Castor, Microtus, Ondatra*) out of thirteen are represented by the same skeletal part in both (sub)assemblages. That two of these four particular genera are the ones represented by the same skeletal part regardless of side is, in this case, easy to explain. Only skulls and mandibles of *Microtus* were identified among the mammal remains at Cathlapotle; postcranial remains of this genus were not identified and this markedly increases the probability that the same skeletal part (regardless of side) will be identified in both components. The

Table 2.11. *The most abundant skeletal part representing thirteen mammalian genera in two (sub)assemblages at Cathlapotle. When more than one skeletal part represented the same MNI, all skeletal parts are listed. R, right; L, left*

Taxon	Precontact assemblage	Postcontact assemblage
Lepus	R proximal tibia	R mandible
Aplodontia	L mandible	R mandible
Castor	L femur	R mandible, R femur
Microtus	L mandible	R mandible
Ondatra	L mandible	R mandible
Canis	R P4	L dP4
Ursus	L m2	R ulna
Procyon	R proximal radius	L m1
Mustela	R mandible	R distal humerus
Lutra	R proximal radius	R mandible, R distal humerus
Phoca	L distal humerus	R temporal
Cervus	L astragalus	L naviculo cuboid
Odocoileus	R calcaneum	R m3, R astragalus

mandible of *Aplodontia* is the most frequent skeletal part in both (sub)assemblages because it was selectively retained by site occupants as a wood-working tool – a chisel or engraver (Lyman and Zehr 2003). It is unclear why the same skeletal part provides the MNImax of *Castor* and *Ondatra* in both subassemblages, but those parts are particularly robust and thus relatively immune to taphonomic attritional processes.

The most frequent skeletal parts in Table 2.11 are from all parts of the skeleton – the head (upper and lower teeth, mandibles), the forelimb, and the hindlimb. It is likely, given what we know about taphonomy at this time, that this is the pattern that will emerge in most cases. Because taphonomic processes influencing the survival and distribution of faunal remains are *not* perfectly correlated with remains that are (or those that are not) taxonomically identifiable (Lyman 1994c), it is unlikely that we will find cases in which the MNImin and MNImax values for a given collection will be perfectly correlated at a ratio scale. They *might* be correlated at an ordinal scale, but it is quite likely even then that the correlation coefficient will be less than 1.0. Shifts in taxonomic abundances will likely not be uniform across all taxa when one shifts from MNImin to MNImax.

Different aggregates of faunal materials making up a total collection will not only produce different MNI values, but they will do so differentially across taxa. Let's say we have one taxon represented in a collection from a site, and that this taxon is represented by twenty-five left and thirty right distal humeri, the most common

Table 2.12. *Fictional data showing how the distribution of most abundant skeletal elements of one taxon can influence MNI across different aggregates. If stratigraphic boundaries are ignored, a minimum of thirty individuals is represented by thirty right humeri. Using stratigraphic boundaries to define aggregates, the total MNI is forty-seven because the most abundant element is left humeri in stratum 1, but the most abundant element in strata 2 and 3 is left humeri*

	Left humeri	Right humeri	MNI per stratum
Stratum 1	22	5	22
Stratum 2	3	17	17
Stratum 3	0	8	8
∑MNI	25	30	47

skeletal part. Obviously, we have a MNImin of thirty (assuming that we find matches in terms of age, sex, and size for all possible pairs of elements; i.e., the twenty-five left specimens all have matching right specimens). But there are also three strata (or horizontally distinct recovery contexts, if you prefer) comprising the site, and the humeri are distributed across those strata as indicated in Table 2.12. When we sum the MNImax values in Table 2.12, we have a site total of ∑MNI = 47. Why? Because whereas with MNImin we had only one most common skeletal part in the form of the right distal humeri (∑ = 30), we now have in Stratum 1 the left humerus as the most common part whereas in Strata 2 and 3 the right humerus is the most common part. The change from one kind of most common skeletal part to two kinds resulted in an increase of 17 MNI (57 percent).

As a final example, let's say we have two taxa. Taxon 1 is represented by the remains of 7 individuals (= MNImin); those remains consist of 7 R humeri, 6 L humeri, 6 R femora, and 5 L femora (∑NISP = 24). Taxon 2 is represented by the remains of 14 individuals (= MNImin); those remains consist of 14 R humeri, 7 L humeri, 6 R femora, and 10 L femora (∑NISP = 37). If we define faunal assemblages stratigraphically, and there are three strata in the site, we may find the stratigraphic distribution of skeletal parts indicated in Table 2.13. In that table the MNI for taxon 1 shifts from MNImin = 7 to MNImax = 10, and the MNI for taxon 2 does not shift but rather both MNImin = 14 and MNImax = 14. The change in taxon 1 is the result of changes in the number of most abundant skeletal parts defined for this taxon as the aggregates change. Most disconcerting is the fact that the ratio of taxon 1 to taxon 2 changes from 7:14 (or 1:2) to 12:14 (or 1:1.2) with a simple change in aggregates. Again, these changes result from specification of different most common skeletal parts with each different set of aggregates.

Table 2.13. *Fictional data showing how the distribution of skeletal elements of two taxa across different aggregates can influence MNI. If stratigraphic boundaries are ignored, there are only seven individuals (R humeri) of taxon 1, and fourteen individuals (R humeri) of taxon 2. Aggregates defined by stratigraphic boundaries produce twelve individuals of taxon 1 and fourteen individuals of taxon 2*

	Taxon 1	Taxon 2
Stratum 1	6 R humeri, 2 L humeri, 3 R femora, 3 L femora (MNImax = 6)	4 R humeri, 1 L humerus, 4 L femur (MNImax = 4)
Stratum 2	1 R humerus, 4 L humerus, 1 R femora, 1 L femur (MNImax = 4)	4 R humeri, 1 L humeri, 3 R femora, 1 L femur (MNImax = 4)
Stratum 3	1 L femur, 2 R femora (MNImax = 2)	6 R humeri, 5 L humeri, 4 R femora, 4 L femora (MNImax = 6)

Changes like those documented above are likely to occur more often than not. This renders MNI a very unstable measurement unit. Of course, changes in aggregation may not cause MNI values to fluctuate *if* the distribution of most abundant skeletal parts is the same for each taxon. Consider, for example, the two-taxon fictional data given in the earlier example, but this time with similar distributions of most abundant elements across the three strata, as shown in Table 2.14. The most abundant element of both taxa (R humerus) has a similar distribution across all three strata and displays its greatest frequency in Stratum 1. The ratio of taxon 1 to taxon 2 is 1:2 in all three strata by the MNImax method. The ratio of 1:2 is given by the MNImin method as well.

Studying the distribution of most abundant elements per taxon across different aggregates may reveal much about site formation and taphonomic history, as Grayson (1979, 1984) noted years ago. I am, however, unaware of any such studies in the literature. This is surprising given interest in *site structure* (e.g., O'Connell 1987). Perhaps the lack of such studies is an instance of benign neglect. Whatever the case, consideration of how aggregation influences MNI deserves more study than it has received because of the insight it will provide to MNI as a measure of taxonomic abundance and also because of the insights it may grant to site structure. In such studies, an aggregate of any kind might be defined – by a site as a whole, by a stratum, by an archaeological feature (each pit, house floor, hearth, etc.), by an arbitrary excavation unit (say, 2 m × 2 m × 10 cm thick), or some combination thereof.

Table 2.14. *Fictional data showing that identical distributions of most common skeletal elements of two taxa across different aggregates will not influence MNI. Note that the NISP per skeletal element is the same as in Table 2.13. Note also that the four distinct skeletal elements have similar frequency distributions across the three strata, and that the ratio of taxon 1 to taxon 2 is 1:2 in each of the three strata, and that MNI values determined while ignoring stratigraphic boundaries also produce a ratio of 1:2*

	Taxon 1	Taxon 2
Stratum 1	5 R humeri, 4 L humeri, 4 R femora, 3 L femora (MNImax = 5)	10 R humeri, 5 L humerus, 4 R femur, 8 L femur (MNImax = 10)
Stratum 2	1 R humerus, 1 L humerus, 1 R femora, 1 L femur (MNImax = 1)	2 R humeri, 1 L humeri, 1 R femora, 1 L femur (MNImax = 2)
Stratum 3	1 R humerus, 1 L humerus, 1 R femur, 1 L femora (MNImax = 1)	2 R humeri, 1 L humeri, 1 R femora, 1 L femora (MNImax = 2)

A final point to consider involves Uerpmann's (1973:311) observation that the "difference between number of finds [NISP] and 'minimum number of individuals' increases as the size of the sample increases" (Uerpmann 1973:311). The difference between NISP per taxon and MNI per taxon will increase as NISP increases (Figure 2.4). Because larger sample sizes allow greater differences between values, differences between MNImin and MNImax will be greatest in large samples and smallest in small samples. Thus, large sample sizes, which are desired for statistical reasons (large samples tend to be more representative than small samples of the population from which they are drawn, and they tend to increase the statistical power of a test), tend to be the ones in which MNI fluctuates the most as different aggregates are defined. As Grayson (1979:210) noted, "This is not the usual behavior of a unit of measurement."

Restating problem seven, MNI measures not only taxonomic abundances but aggregation methods as well (Grayson 1984). This can be shown graphically and statistically by considering the relationship between NISP and MNI as modeled in Figure 2.4, but with log-transformed data such that the relationship is linear. The slope of the simple best-fit regression line describing the relationship between NISP data and taxonomically corresponding MNI data should be less steep when more agglomerative (larger aggregates) methods are used to calculate MNI (MNImin) than when less agglomerative (smaller aggregates) methods are used to calculate

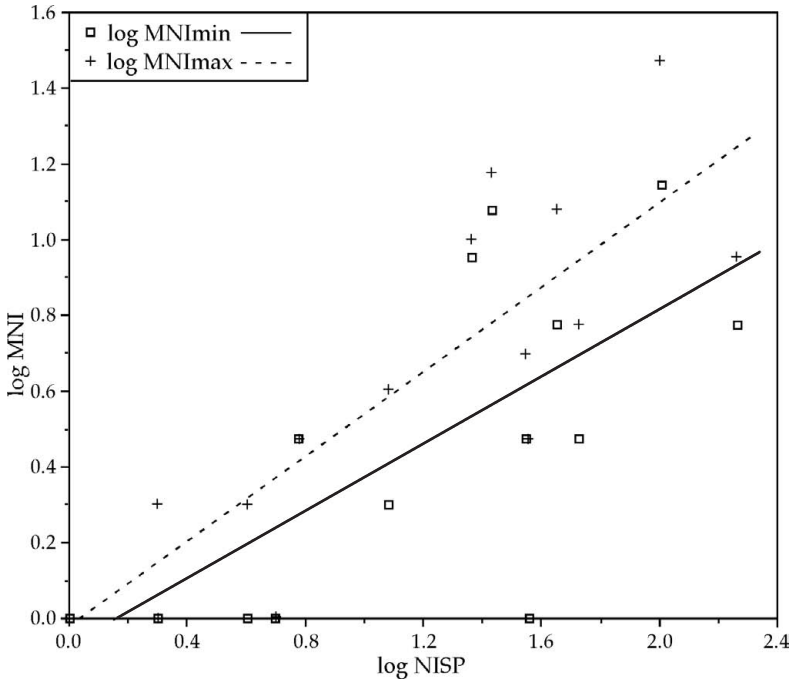


FIGURE 2.11. Relationships between NISP and MNImin, and NISP and MNImax at site 45DO214.

MNI (MNImax). This is so because MNImin involves few most common elements so it is difficult to find and thus add a new most common element. MNImax involves many most common elements so it is easy (relatively speaking) to find and thus add a new most common element. Consider site 45DO214 among the collections from eastern Washington State mentioned previously (Table 2.7). The slope of the simple best-fit regression line describing the relationship between NISP and MNImin is 0.44 (Table 2.7); the slope of the simple best-fit regression line describing the relationship between NISP and MNImax is 0.57. Both sets of data points and both best-fit regression lines are included in Figure 2.11, which shows that the slope for MNImin is less steep than that for MNImax.

The example of 45DO214 indicates that MNI not only measures taxonomic abundances but it also measures aggregation. But we do not want measures of two variables to obscure each other, particularly when one may approximate the target variable and the other has nothing whatsoever to do with the target variable. Is there, perhaps, a logical way to define aggregates such that the measures of taxonomic abundances provided by MNI can be treated as if the aggregates do not significantly influence those abundances?

Defining Aggregates

Recall that an *assemblage* of faunal remains is the set of remains from a horizontally and vertically bounded space, usually a geological space such as a stratum or a part thereof. Following Grayson (1984), the term *aggregate* is used as a synonym for assemblage, and the term *aggregation* for the process of defining the spatial boundaries of a faunal assemblage. Despite Grayson's (1973) recognition of the aggregation problem more than 30 years ago, few analysts other than Grayson (1979, 1984) have subsequently explored its implications with their own data. Thus it is not unusual to find paleozoologists still calculating MNI values without considering the aggregation problem (e.g., Trapani et al. 2006). The aggregation problem is not even mentioned in one textbook on zooarchaeology (Rackham 1994).

Payne (1972) suggests that aggregates of faunal remains should be defined on the basis of the homogeneity of taxa and their frequencies. Ignore for the moment the question of how similar is similar enough for two assemblages to be considered homogeneous (Payne did not address this question), and consider the following three things. First, this procedure assumes that *natural* faunal aggregates exist and we have but to discover them. But whether natural discoverable faunal aggregates exist or not is unclear. Furthermore, what kinds of faunal aggregates are to be searched for – those representing a depositional event, a human-behavior, a death event, or . . . what? Perhaps the research question being asked would help specify the appropriate aggregate, but other problems attend their definition. The second thing to consider, then, is that Payne's protocol precludes the study of stasis because one aggregates, say, stratigraphically sequent, similar (homogeneous) assemblages. Finally, Payne's procedure comprises a circular process: Change in the fauna would be identified based on how the aggregates were defined – the property of differences or nonhomogeneity interpreted as change – because aggregates are defined on the basis of similarity and homogeneity. Based on these three observations, we might use nonfaunal criteria to define aggregates.

Ringrose (1993:128) argues that "not all levels of aggregation are likely to be sensible, so that the problem [of the influence of aggregation on MNI] is perhaps less than it might seem at first. If it is not possible for specimens from the same individual to be present in two locations [that is, to be tallied in two distinct aggregates], then it is nonsensical to calculate the MNI at a level of aggregation where these two locations are taken together, since specimens will be, implicitly, counted as being possibly from the same individual when in fact they cannot be." Implementing this protocol of defining aggregates demands a great deal of knowledge regarding the taphonomic history of the materials under study. Some of it might be found by refitting studies

(e.g., Rapson and Todd 1992; Todd and Stanford 1992). However, this is again using faunal data to define aggregates and thus imparts a degree of circularity to those definitions. It also can introduce the problem of matching potentially paired (left and right) skeletal parts (e.g., Todd and Frison 1992).

Some zooarchaeologists indicate that one should define faunal aggregates based on "cultural units rather than arbitrary ones related to excavation logistics" (Reitz and Wing 1999:198). This is all well and good, but does this mean that two pits containing bones and originating in the same stratum (dating to the same time period and apparently representing the same cultural context given stratigraphic contemporaneity) should be considered separately or together? Archaeologist James Ford (1962) argued long ago that archaeological "cultural units" such as cultures, phases, periods, and the like were often defined on the basis of stratigraphically bounded aggregates of artifacts, but that it was unclear why there should be any necessary relationship between sediment deposition boundaries and boundaries between cultures. I agree (Lyman and O'Brien 1999). So what are we to do?

Let us begin by glancing at the solution that paleontologists have used. Fagerstrom (1964:1198), a paleontologist interested in past biological communities, suggested that a *fossil assemblage* representing a community was "any group of fossils from a suitably restricted stratigraphic interval and geographic locality." What is suitable is not clear, though it is hinted at in other paleontological concepts. In vertebrate paleontology, a *faunule* is an assemblage of associated animal remains from one or several contiguous strata, dominated by members of one biological community (Tedford 1970:677). And, a *local fauna* is a set of remains from one locality or several closely spaced localities which are stratigraphically equivalent or nearly so, thus it is a set of taxa close in (geological) time and (geographic) space (Tedford 1970:678). Identifying prehistoric faunal communities – or faunules – was what Shotwell (1955, 1958) was concerned about, and he emphasized the taphonomic problems with doing so when one used an aggregate of fossils the boundaries of which were set by excavation strategies and stratigraphy.

The preceding brief discussion hints at two things. First, paleontologists often seek, like Chester Stock and Hildegard Howard did, to determine the census of a paleocommunity, or a biocoenose. That is their target variable, and they acknowledge the geological mode of occurrence of the faunal materials that they study, and they use stratigraphic boundaries and extent of exposures to collect a sample of those materials. The second thing hinted at is an extremely critical detail. Reitz and Wing (1999:197) mention it when they state that the aggregates of faunal remains defined "may depend on the research problem." Any aggregates defined *must* depend on the research problem, as well as whatever taphonomic and site-formational information is available. Thus, on the one hand, human behaviorally significant assemblages of

remains, such as those in cache pits or in trash middens or on house floors are likely to be important to questions about human interactions with fauna. On the other hand, questions regarding paleoecology are likely to be phrased in such a manner as to require temporally and spatially distinct assemblages of remains, perhaps but not necessarily representing one "biological community" but certainly providing insights to the nature of biocoenoses. Temporally distinct assemblages but perhaps not human behaviorally significant ones would be of interest to paleoecologists.

Valensi (2000:358) noted that aggregation based on excavation levels "gave an overestimation of MNI [as a result of specimen] interdependence [across] some levels." Interdependence was recognized by refitting specimens of both lithic and bone specimens that came from different depositional units. Valensi used archaeostratigraphic units as the basis for defining aggregates, and found refitting specimens that came from different units. Her analysis suggests a protocol for defining aggregates. Refits of lithic specimens would provide nonfaunal criteria for defining faunal aggregates. The paleozoologist could adopt a rule, such as only when refits across aggregates are minimal, whereas refits within aggregates are maximized, have appropriate aggregates for determining MNI been defined. However, not only is the time cost incredibly high if the assemblage is large – do the faunal refits, using the same rule, define the same aggregates as the lithics (or ceramics)?

Research questions about taphonomic histories likely will require an estimate of a taphocoenose, those about hunter or predator selectivity will require not only an estimate of a thanatocoenose but also the biocoenose from which it derived. Explicit statement of the research problem and research questions should help the paleozoologist define aggregates that are pertinent. Of course, any available taphonomic information such as obvious refits should also be consulted to help set geological spatial boundaries around the aggregate(s). This does not mean that one will automatically have aggregates that do not share specimens from the same individual, but perhaps those will be so rare as to not significantly bias any statistical results.

DISCUSSION

Thus far the problems with NISP as a quantitative unit giving valid measures of taxonomic abundances (even in a taphocoenose, let alone in a thanatocoenose or biocoenose) have been considered and it has been argued that all but one of those problems – that of possible specimen interdependence – can be fairly easily resolved analytically. (Some analysts still fail to realize how easily many of the problems with NISP can be resolved analytically [e.g., O'Connor 2001].) Problems with MNI as a quantitative unit giving valid measures of taxonomic abundances have also been

identified and discussed, and it has been shown that many of those are also readily dealt with analytically. The problem that remains with MNI is aggregation. As implied above, there is no magic algorithm for solving the aggregation problem because each aggregate specified by the analyst may, or may not, have a set of faunal remains all of which are indeed independent of all other faunal remains in all other aggregates. Earlier I referred to the latter as *Adams's dilemma*. It is aptly referred to as a dilemma because if, say, stratigraphically bounded aggregates are chosen as the assemblages to be analyzed, one must assume that the faunal remains in each are independent of all other faunal remains in other aggregates. But, of course, they might not be.

A chosen sampling design may indicate where to excavate and which screen-mesh size to use, but the faunal specimens recovered are a result of the taphonomic history of the assemblage – which bones and teeth were accumulated, deposited, and still exist, and where they are located, both horizontally and vertically. The existing remains of a single individual may be in one or more horizontal loci, in one or more vertical loci (or strata), or both. (No two specimens can, of course, occupy exactly the same horizontal and vertical position. By *same location*, I mean a spatially limited, horizontally and vertically bounded unit.) Even attempting to match and pair all skeletal specimens from all excavated recovery proveniences, we will likely never know what the *correct* aggregates of faunal remains should be. By correct is meant those that are not only relevant to our research questions, but also ones defined such that specimens from a single carcass are *not* distributed across two or more aggregates. Given that we cannot know this, we either assume Adams's dilemma does not exist, or, we do something other than determine MNI values.

There is, in fact, a relatively simple solution to Adams's dilemma. The solution rests on the fact that quite often, virtually the same information regarding taxonomic abundances in an assemblage is found in NISP as is found in MNI. This statistical relationship has been known for some time (Casteel 1977, n.d.; Grayson 1978a, 1979). In short, MNI is redundant with NISP, where "redundant" means that the two quantitative units produce the same information. The "same information" can mean identical, or simply statistically indistinguishable. To show that MNI and NISP provide the same information in both of these senses, consider the owl pellet data mentioned before. Recall that the sample comprises eighty-four pellets, that the relationship between NISP and MNImin is linear (Figure 2.8), and that the relationship is strong ($r = 0.989$, $p < 0.0002$). For this sample, 97.8 percent of the variation in MNI values is explained by variation in NISP values. Clearly, MNI is redundant with NISP. And, the same applies to the fourteen samples of mammal remains from eastern Washington State (Table 2.7). For these assemblages, the relationship between NISP and MNImin is typically strong ($r > 0.75$ for 13 of the 14) and significant ($p < 0.01$ for all). For thirteen of these fourteen assemblages, NISP accounts for

more than 51 percent of the variation in MNI. MNI provides information about taxonomic abundances that is redundant with that provided by NISP (Figure 2.4). But so what? Constructing an answer to this question requires a consideration of the scale of measurement represented by NISP and by MNI.

WHICH SCALE OF MEASUREMENT?

Some years ago, Grayson (1984:94–96) noted several critical things. First, he noted that converting from one ratio scale to another ratio scale based on different measurement units will not alter the value of a ratio of measurements. This is so because both ratio scales have natural zero points and their respective units of measurement stay constant in each. Thus, the ratio of the weight of two items will not alter if first measured in pounds and then in kilograms. If the two items are 50 pounds and 75 pounds, the ratio of their weights is 1:1.5; the two items weigh 22.68 kilograms and 34.02 kilograms, respectively, for a ratio of 1:1.5. As noted earlier in this chapter, aggregation has the unsavory characteristic of altering MNI tallies, thereby causing ratios of taxa to change as the manner in which faunal remains are aggregated changes.

The second thing Grayson (1984) noted was that MNI values are not ratio scale values precisely because they are *minimum* numbers (Table 2.4). The *actual* number of animals represented (by the identified assemblage) could be as great as the NISP, although it likely will fall somewhere between the MNI and the NISP given the probability (> 0.0) of some interdependence of specimens. Thus it cannot be argued that a taxon represented by an MNI of ten is half as abundant as a taxon represented by an MNI of twenty, nor can it be argued that if two taxa each have MNI values of fifteen they are equally abundant. Figure 2.12 plots ratios of the abundances of each pair of taxa based on NISP, MNImin, and MNImax measures of the four least common taxa in the collection of eighty-four owl pellets (Table 2.9). The ratios vary by greater or lesser amounts across the three quantitative measures. In particular, note the variation in ratios between the MNImin and MNImax values. There is no way to determine which set of ratios most closely measures the actual abundances of taxa. Clearly, it is ill-advised to treat MNI values as ratio scale because there are many reasons why they likely are not.

Unfortunately, it is unlikely that NISP values are ratio scale. As Grayson (1984) noted, if MNI provides *minimum* tallies, NISP provides maximum tallies. Given that we do not know the nature (extent) of interdependence of the specimens comprising the NISP for any given taxon in any given collection, and given that intertaxonomic variation in such interdependence will differentially influence how closely NISP tracks a taxon's actual abundance, it is unlikely that ratios of taxa based on NISP values

Table 2.15. *Ratios of abundances of pairs of taxa in eighty-four owl pellets. Original data from Table 2.8. Taxon 1,* Sylvilagus*; taxon 2,* Reithrodontomys*; taxon 3,* Sorex*; taxon 4,* Thomomys

Taxon pair	NISP	MNImin	MNImax
1–2	0.26	0.20	0.40
1–3	0.11	0.20	0.29
1–4	0.07	0.12	0.17
2–3	0.41	1.00	0.71
2–4	0.28	0.62	0.42
3–4	0.68	0.62	0.58

are in fact ratio scale. There is no way to know which set of ratios of abundances of taxa in the owl pellet fauna (Table 2.15), if any, most accurately reflects the actual ratio scale abundances of the taxa. As Grayson (1984:96) noted, because "we know nothing of the nature of the frequency distribution [of taxonomic abundances] that begins with MNI and ends at NISP for a set of taxa," knowledge of ratio scale abundances of taxa is precluded.

If MNI and NISP do not provide ratio scale taxonomic abundance data, do they perhaps provide ordinal scale abundance data? Again, Grayson (1984:96–99) provided

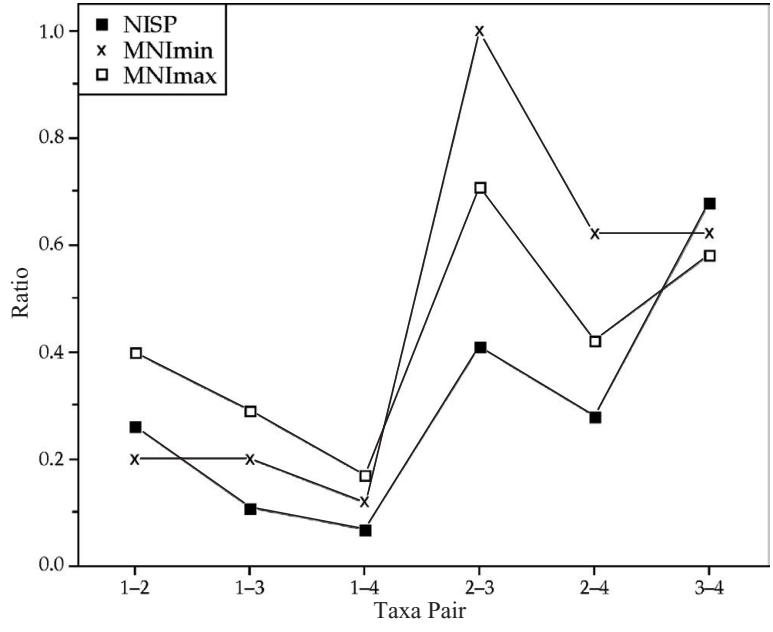


FIGURE 2.12. Ratios of abundances of four least common taxa in a collection of eighty-four owl pellets based on NISP, MNImax, and MNImin. Taxon 1, *Sylvilagus*; 2, *Reithrodontomys*; 3, *Sorex*; 4, *Thomomys*. Data from Table 2.8.

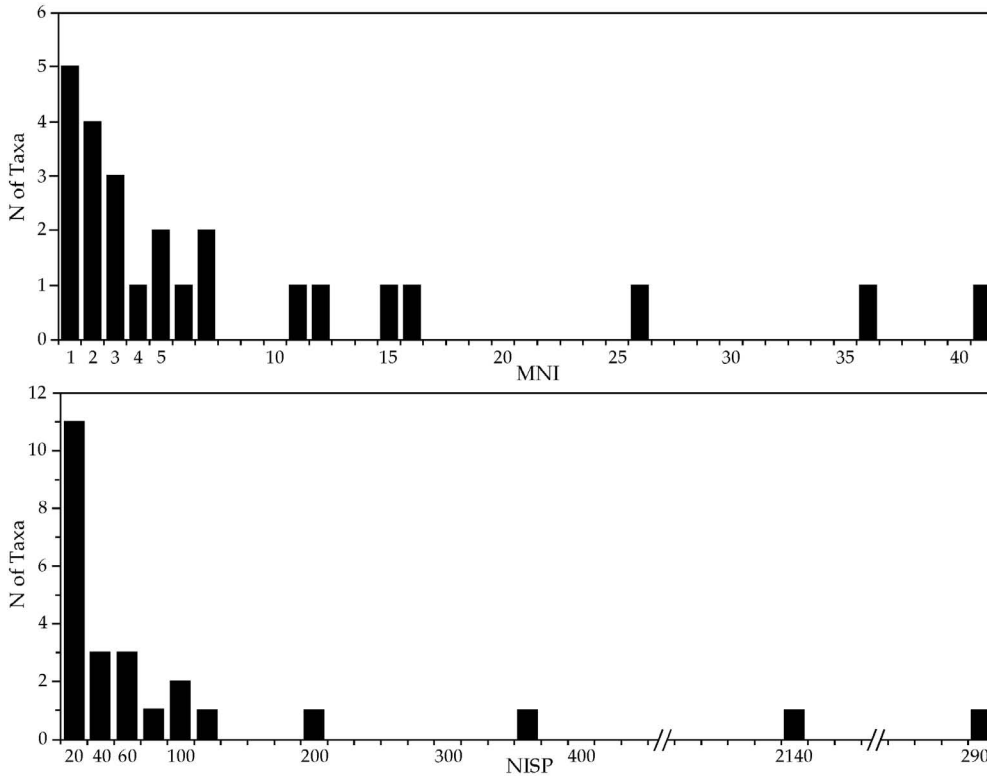


FIGURE 2.13. Frequency distributions of NISP and MNI taxonomic abundances in the Cathlapotle fauna. Data from Table 1.3.

a clear answer. The rank order abundances of taxa are often quite similar across NISP and MNI; if the two sets of values are significantly correlated on an ordinal scale, then the included taxonomic abundances are ordinal scale. Why NISP and MNI should often be correlated comprises the critical insight as to why we can conclude they are ordinal scale. In most multitaxa faunas, a few taxa are represented by many individuals and specimens, and many taxa are represented by few individuals and specimens. As taxonomic abundances increase (whether NISP or MNI), the magnitude of the differences between abundances of adjacent taxa increases. Such frequency distributions increase the probability that taxonomic abundances are ordinal scale because there is less chance that variation in aggregation (MNI) or specimen interdependence (NISP) will alter rank order abundances.

Summing the precontact and postcontact assemblages, eighteen taxa in the Cathlapotle fauna (Table 1.3) are represented by seven or fewer individuals whereas only seven taxa are represented by more than ten individuals (Figure 2.13). Similarly, 20 taxa are represented by 100 or fewer specimens whereas only 5 taxa are represented by