

Bucknell University Interlibrary Loan



ILLiad TN: 97081

**Borrower:** GDC

**Lending String:** \*PBU,PBB,FWR,TNY,FWR

**Patron:** ;dept; ;type; Amith, Jonathan

**Journal Title:** Language in society.

**Volume:** 8 **Issue:**  
**Month/Year:** 1979 **Pages:** 337-64

**Article Author:**

**Article Title:** Ellen, Roy F; Omniscience and  
ignorance

**Imprint:** Cambridge, Eng. [etc.] New York,  
Cambrid

**ILL Number:** 13384769



**Call #:**

**Location:**

**ARIEL**  
**Charge**  
**Maxcost:** \$25IFM

**Shipping Address:**  
Gettysburg College Library  
300 N. Washington St.  
InterlibraryLoan  
Gettysburg, PA 17325-1493  
IDS # 132  
**Fax:** (717) 337-7001  
**Ariel:** 138.234.152.5

**Omniscience and ignorance: Variation in  
Nuauulu knowledge, identification  
and classification of animals**

ROY F. ELLEN

*University of Kent at Canterbury*

(With an appendix by David Reason of the University of Kent at Canterbury)

ABSTRACT

The ethnographic analysis of categories is still largely based on assumptions of cultural uniformity, although, during the past decade, the significance of variation has become increasingly evident as attempts have been made to measure it. Delineation and measurement are themselves complex tasks, however. In a single body of data there may be variation according to many criteria which are often cross-cutting and reinforce each other irregularly. These issues are discussed in this paper in relation to different types and contexts of variation evident in animal classifications of the Nuauulu of eastern Indonesia. Yet, the kinds of assumptions made in formal studies of individual variation are as problematic as those concerning cultural uniformity. It is important to appreciate that the techniques and representations employed to describe classifications and their variation are often inadequate, concealing those things that are operationally of most significance and reifying 'classifications' which do not always exist in practice. The products of classifying behaviour inevitably reflect the immediate social conditions of the situations in which they are used. (Analysis of categories, cultural variability, ethnozoology, social context; Nuauulu of eastern Indonesia.)

INTRODUCTION

Despite some determined efforts to eliminate him, most published reports on ethnobiology still list folk categories and classificatory constructs on the basis of the hypothetical omniscient 'speaker-hearer'. Much the same is true for the ethnographic analysis of any kind of category. This non-existent figure has the distinction of possessing a maximum knowledge of the corpus under examination, the sum total of the knowledges of all persons consulted, without reflecting the practical (and generally partial) knowledge of individuals or indicating the dimensions and degree of variation. Yet if data are assembled using this heuristic device, they will always be in excess of those pragmatic competences common to

all individuals. Moreover, they will not reveal unique abilities or those shared by only part of the population. These are matters which may be of considerable sociological significance. So, since there is a large area with an uneven range of abilities, and since no techniques of elicitation are ever completely exhaustive, variation must be expected from any corpus, even if the data provide no means of revealing it.

There are a few early references in the ethnobiological literature to cognitive variation<sup>1</sup> between individuals and between groups (e.g. Henderson & Harrington 1914: 8). In some recent publications, it is recognized as a problem area (Gal 1973; Berlin, Breedlove & Raven 1974: 58-9), but there have been only limited attempts to tackle it as an integral part of the theory and structure of classification systems (cf. Wallace 1970). In the past, this was partly due to assumptions about the shared nature of culture, but also to the small size of the samples involved. With the larger samples of more recent work patterned variation is inevitably revealed. However, little quantitative information has been presented on the distribution of response variability in interpretations between informants according to such normally important variables as geography, age, gender, kinship affiliation, ideology, degree of literacy, and so on (Hunn 1975: 16-18, 21; Manning & Fabrega 1976: 41-3). Two notable exceptions are found in the work of Gardner (1976) and Hays (1974, 1976).

The remarks which follow arise from ethnozoological inquiries over a number of years among the Nuaulu people of South Seram, eastern Indonesia.<sup>2</sup> Although these inquiries extend to most major groups of animals, the illustrations used here are limited, having been largely drawn from an analysis of categories for amphibia and reptiles (Ellen, Stimson & Menzies, 1976, 1977, 1977a). A more complete monograph is currently in preparation. Part of this work has been directly concerned with mapping and measuring variation in classifications used by individuals, how this varies according to context, and how absolute zoological knowledge of a particular domain varies within a community. It has sometimes

[1] Most cognitive variation is necessarily expressed lexically, but different words do not always indicate categorical differences. I am not concerned here with alternative labelling which can be explained through synonymy, dialect or other linguistic usages (Gardner 1976: 447-52; Hage & Hawkes 1975: 299; and works cited).

[2] Fieldwork was undertaken from 1969 to 1971 and in 1973 and 1975, supported by research grants from the British Social Science Research Council, the Central Research Fund of the University of London, the London-Cornell Scheme for East and Southeast Asia and the Hayter Travel Awards Fund. This paper is a direct outcome of an SSRC award for research on 'Nuaulu ethnobiology and ecology' covering the period 1975-7, sponsored by the Indonesian Institute of Sciences and the Museum Zoologicum Bogoriense. I have acknowledged elsewhere (Ellen, Stimson & Menzies 1976, 1977, 1977a) my indebtedness to various individuals on specific matters relating to the data discussed here. Once again, I am happy to be able to thank Ralph Bulmer and David Reason for their comments on earlier versions of this paper. Dell Hymes raised some useful points in correspondence.

#### OMNISCIENCE AND IGNORANCE

highlighted an unsuspected degree of variability in such systems (Ellen 1975). Here I discuss some of the problems raised in measuring variation, and survey its possible causes in terms of differences between individuals, contexts, and the properties of the various kinds of natural species involved, as well as interference from the methods employed. I argue that this provides the means of placing the study of classification systems in a firmer sociological context, allowing us to see how arrangements of categories relate to concrete usage situations, rather than relegating their analysis to broad speculation about mental principles of cognition.

#### THE MEASUREMENT OF VARIATION

Given the existence of variation, it is necessary to find some standard from which it can be measured. With biological taxonomy in the scientific community, this standard is embodied in a widely agreed-upon written record and set of conventions. It has authority which stems from a recognized pedigree, general convenience, and a respect for precedence. Most importantly, it is constructed around evolutionary theory, its units being at least putatively genetic or phylogenetic groupings. Prescriptive codifications meticulously determine nomenclatural usage.

For the folk-classification of many pre-capitalist societies there is no written record. Authority may come from persons respected for their knowledge of certain matters or their ascription or achievement in other fields; it may come from oral tradition and normatively accepted notions as to the order of the natural world. In such situations, contradiction (sometimes self-contradiction), variation, and flexibility must be expected as a normal characteristic of classification systems (Bulmer 1975: 24; Harris 1968: 586; Tyler 1969: 5), and the premise of a 'correct' or 'true' identification must be more difficult to establish in fact (Bloch 1971: 86). As Bulmer (1969: 6) comments:

The terms used to denote the units of a folk taxonomy are part of natural language, and as such can be used intuitively, imaginatively, with different referents in different contexts, in contrast to terms used in biological taxonomy, usage of which is explicitly restricted with the object of minimizing ambiguity.

Part of the problem facing anyone concerned with making sense of this kind of classification is precisely that many analysts have *a priori* attributed to them the characteristics of the scientific, phylogenetic scheme of things. We cannot assume that the kind of logic employed in folk-classification is simply discovered through the distinctive-feature analysis (usually based on binary discriminations) developed by post-Linnaean taxonomists. 'If folk taxonomies could be satisfactorily analysed in terms of straightforward series of binary oppositions applied to single characters,' writes Bulmer, 'the most economical way of glossing terms

applied to these would be to codify . . . [them], and to that extent zoological identifications of the creatures classified would be redundant' (*ibid.*: 4).

The analogy of phylogenetic taxonomy has led to the concealment of variation and a blindness to the idea that definitive attributes may, in fact, keep changing. One individual may see *p*, *q* and *r* as the distinctive characteristics of  $\Delta$ , another may select *r*, *s* and *t*, and the same person, on another occasion, may decide on *t*, *u* and *v*. Here no single feature is essential for membership of a category, nor is any such feature sufficient for membership. It is very likely that such polythetic processes explain a considerable amount of the apparent variation in folk classifications (Needham 1975: 349-50; Ellen 1979), although it by no means explains why particular features are chosen.

The polythetic model, rather than conceiving of categories in terms of the presence or absence of particular features (which may be appropriate in any single elicitory context), sees them as configurations of features, where emphasis on a particular one varies according to context or individual. Whether anomalies are endemic or exceptional therefore becomes a question of circumstances, and particular anomalies may depend very largely on the context of elicitation (Ellen 1975: 211-18, 221). This notion of a configuration of features has led some commentators to talk in terms of *mental images* (Simpson 1961: 12; Bulmer & Tyler 1968; Hunn 1973), but such a concept seems altogether *too* holistic and rigid, given some of the recent empirical and experimental evidence. *Gestalt* is better (cf. Hunn 1975: 10-13), but we might do well to remember that, in this case, we are dealing with *gestalts* that are by no means fixed and often vary according to context.

Physical nature may well condition and constrain classification and on occasion the degree of consensus might be quite remarkable, but it is of great importance to stress variation between individuals and collectivities in their knowledge of actual species, of lists of names and categories, and the ways they arrange these in classifications, as well as variation in knowledge and classification through time and according to changing contexts (Bricker 1974; Ellen 1975, 1976: 6; Friedberg 1974; Harris 1970: 12; Hays 1974 (and works cited, p. 359, footnote 8); Manning & Fabrega 1976: 42-3 (and works cited); Rosaldo 1972: 83-4; Sankoff 1971). The greatest problem so far has been to record and work out the degree and significance of elasticity. In domains where the dimensions of variation are limited, a greater degree of success has been achieved (Heider 1972; Hage & Hawkes 1975). For example, categories of colour may be, and often are, abstract summary terms or based on a very particular aspect of what is observed. At least, the ethnographer with Munsell colour chips assumes this to be the case. The dimensions of variation are therefore artificially limited. The problems attached to the measurement and understanding of variation for the complex categories of natural species are much more formidable.

CONSISTENCY, SHARING AND FLEXIBILITY

In measuring cognitive variation we are generally dealing with three variables: lexical items (*L*), the content of categories in terms of scientific taxa (*X*), and informants (*I*). Others, such as knowledge of uses, could be added, but as yet this is an aspect which has not been given wide attention (but see Hays 1974: 305-39). Moreover, although the non-congruence between words and categories is well understood, it is generally assumed that lack of congruity and the existence of covert categories can, in most cases, be ignored, and that they inevitably must be so if measurement is to be a possibility. Given this, we can go on to distinguish three basic kinds of variation:

- (a) variation in the relationship between lexical items and scientific taxa;
- (b) variation between informants; and
- (c) variation according to rules appropriate to specific contexts.

These can be described, respectively, as *consistency*, *sharing*, and *flexibility*. However, in discussing the dimensions of variation, it is convenient and illuminating to also distinguish between *identification* and *classification*. In one sense, there is no difference between the two: both concern the allocation of an item to a (usually) named category. A distinction can be drawn, however, in terms of the way in which informants assign observed specimens to terminal categories and the way in which categories are arranged in more inclusive groups, i.e. the difference between the statements 'x (an object) is a y (category)' and 'y is a kind of z'. Combining these two sets of distinctions, we get the following matrix:

	consistency	sharing	flexibility
identification	1	3	5
classification	2	4	6

Variation between lexical items and scientific taxa is about *consistency in identification*. Thus, the majority of a series of Nuauulu informants associated the frog category *kere* with specimens of *Platymantis papuensis* and only one with *Rana modesta*. On the other hand, *R. modesta* was labelled by nine individuals as *notu* and by only one as *kere* (Ellen *et al.* 1976; cf. Hays 1974: 185-6).

One way of measuring this kind of consistency is the calculation of *M*, the proportion of informants opting for the most frequently-cited item, or, alternatively, the proportion of informants opting for the most frequently-cited taxon. However, *M* gives no indication of the diversity of labels employed, and its

accurate use depends on the solution of a number of additional technical difficulties. A different measure,  $\Delta p$ , relates to the element of predictability which the data allows, while  $V$ , a measure of qualitative variation taking into account all features of the distribution of responses, compares the actual extent of disagreement with the maximum conceivable extent.  $\bar{D}$  and  $\hat{D}$  provide measures of deviation from a criterion distribution. In Table 1, these various measures,

TABLE 1. *Some measures of consistency in the application of terms to amphibian and reptile species by Nuaulu informants\**

Amphibia	$n_i, n_j$	$N$	$L$	$M$	$\Delta p$	$V$	$\bar{D}$	$\hat{D}$	No. of specimens
<i>Rana</i> sp.	12	12	1	1.00	0.00	0.00	0.00	0.00	12
<i>Litoria infrafrenata</i>	1,11	12	2	0.92	0.42	0.31	0.70	0.04	12
<i>Rana modesta</i>	9,1	10	2	0.90	0.40	0.36	0.64	0.04	7
<i>Litoria amboinensis</i>	6,1	7	2	0.86	0.36	0.49	0.51	0.08	6
<i>Litoria</i> sp. (bicolor group)	3,7	10	2	0.70	0.20	0.84	0.16	0.36	6
<i>Phrynomantis fusca</i>	2,1	3	2	0.67	0.17	0.89	0.11	0.44	3
<i>Platymantis papuensis</i>	2,6,6	14	3	0.45	0.12	0.92	0.08	0.80	7
Reptilia									
<i>Cuora amboinensis</i>	2	2	1	1.00	0.00	0.00	0.00	0.00	3
<i>Gekko vittatus</i>	1	1	1	1.00	0.00	0.00	0.00	0.00	1
<i>Varanus indicus</i>	7	7	1	1.00	0.00	0.00	0.00	0.00	3
<i>Eugongylus rufescens</i>	1	1	1	1.00	0.00	0.00	0.00	0.00	1
<i>Python reticulatus</i>	5	5	1	1.00	0.00	0.00	0.00	0.00	5
<i>Acanthophis antarcticus</i>	3	3	1	1.00	0.00	0.00	0.00	0.00	1
<i>Dendrelaphis pictus</i> †	8,2	10	2	1.00	0.00	0.00	0.00	0.00	2
<i>Boiga irregularis</i> †	5,3	8	2	1.00	0.00	0.00	0.00	0.00	3
<i>Hemidactylus frenatus</i>	1,50	51	2	0.98	0.48	0.08	0.92	0.00	42
<i>Emoia cyanura</i>	9,1	10	2	0.90	0.40	0.36	0.64	0.04	7
<i>Calotes cristatellus</i>	21,3	24	2	0.88	0.38	0.44	0.56	0.06	14
<i>Candoia carinata</i>	7,1	8	2	0.86	0.36	0.44	0.56	0.06	7
<i>Dasia smaragdina</i>	2,1	3	2	0.66	0.16	0.88	0.12	0.44	2
<i>Mabuya multifasciata</i>	7,17,3	27	3	0.63	0.30	0.78	0.22	0.45	21
<i>Mabuya</i> 'rudis'	2,2	4	2	0.50	0.00	0.00	1.00	1.00	4
<i>Carlia fusca</i> ‡	2,1,5	8	3	0.50	0.17	0.80	0.20	0.44	7
<i>Draco lineatus</i> §	1,1,1,2	5	4	0.40	0.15	0.96	0.04	0.64	5
<i>Emoia kuekenthali</i>	—	—	—	—	—	—	—	—	1

\* Based on 245 informant responses for 182 specimens: see Ellen *et al.* (1976: Table 2; 1977: Table 2; 1977a: Table 2). For explanation of notation see appendix. Species are ordered by zoological class in terms of decreasing  $M$  values.

† *Dendrelaphis pictus* and *Boiga irregularis* are species for which the Nuaulu use different terms for the mature and immature stages. In each case these terms were applied only to specimens of one species, and at this level must therefore be said to be used consistently. However, there is evidence for inconsistency in the use of species-specific terms for the different developmental stages (Ellen *et al.* 1977a).

‡ *Carlia fusca* and *Carlia* sp. are treated as a single species.

§ Refers to identification with the term *hohone* and not sub-specific terminal forms within this category.

which are discussed in more detail by David Reason in his appendix, have been applied to data on Nuaulu identification of amphibians and reptiles, where  $N$  is the total number of responses from different informants and  $L$  the number of different lexical items applied to the taxon.

Clearly, the significance of the measures is very much dependent on the representativeness of the informants. When dealing with such small numbers of informants and specimens, the variable context of the naming task, the variable adequacy of stimuli, 'pseudo-compliance', uncorrected errors in naming and responses of varying degrees of specificity are all going to affect the outcome (Hays 1974: 262-75). For example,  $M$  values of 1.00 indicate a high degree of consistency. This seems probable for common and 'significant' species such as the monitor lizard, *Varanus indicus*. But although many were seen and informally identified, only three were actually collected. On the other hand, some species with values of 1.00 are much rarer (e.g. the death adder, *Acanthophis antarcticus*). It was precisely for this reason that it was impossible to obtain many specimens. Had there been more specimens, there might have been more disagreement. Only larger, more statistically significant samples, uniform numbers of identifications and specimens for each species, and the removal of other interfering variables would allow us to place a general confidence in such measures. Until then, they are at best indicative of tendencies which must be interpreted carefully in the light of the unmeasurable experience of the researcher.

*Consistency in classification* concerns variation in the relationship of non-terminal lexical items (or categories) and their content in terms of scientific taxa. For example, the Nuaulu apply *tekene* to most snakes, but may sometimes include certain species of fish and exclude certain other snakes (Ellen *et al.* 1977a). *Peni*, on the other hand, invariably refers only to cassowary, pig and deer, while *manwe* includes all birds, but sometimes excludes the cassowary (Ellen 1975). Now, in the first case, the variation is partly due to the memorate knowledges of different informants and is perhaps also a function of the artificial elicitation situation. In the second case, there is the additional factor of variation according to rules appropriate to particular social contexts. There is thus overlap between cells 2 and 6 of the matrix.

*Variation between informants* covers measures of the extent of sharing of particular lexical items, lexicon size, numbers of categories, knowledge of scientific taxa, uses and suchlike. It may combine this with the issue of consistency of identification, through a concern with the variation between informants in the way in which they assign scientific taxa to particular lexical items, and the frequency of variation in naming responses. It may combine it with classification in the comparison of profiles of individual taxonomic models (Hays 1974: 237), or in attempts to show how specific classificatory structures vary between individuals. This can be done through the use of tree-diagrams elicited from individual informants (e.g. Coxon 1977). Gardner (1976: 455-7), on the other



hand, has used Venn diagrams to show variation in bird domain shape from the Dene of the Mackenzie drainage area of Canada. Venn diagrams have certain advantages over 'trees', one being that they are more appropriate indicators of 'fuzziness'. I have collected Nuaulu data using card-sorting techniques, which could be represented either way. For example, Fig. 1 illustrates three different ways in which informants classified one portion of the category *poso*, in this sense including all 'lizards'. But because of doubts about the statistical validity of these data, I have preferred to see them simply as an indication of the possible dimensions of variation rather than as any precise measure of distribution.

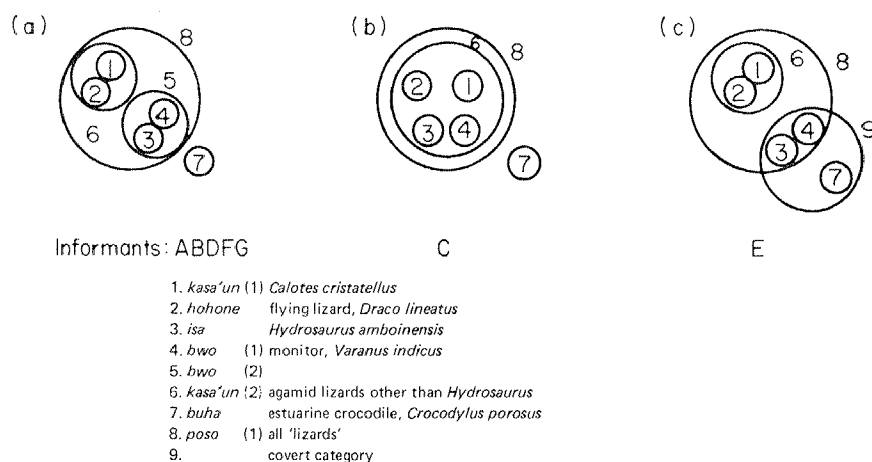


FIGURE 1. Variation between informants in classifying one portion of the Nuaulu category *poso* (lizards). (All identifications in this and subsequent figures are those most frequently provided.)

Furthermore, Venn diagrams, like taxonomies *sensu stricto*, assume that classificatory space is two-dimensional. There are good reasons for portraying classifications and their individual variants as three-dimensional networks (see below), but the technical problems in doing so are formidable. Accordingly, I have tended to incorporate data on variability into general charts represented as taxonomies (Fig. 2; see also Ellen *et al.* 1977, 1977a). These may serve as perfectly convenient indicators of variability without suggesting that they have the accolade of statistical authority.

One problem with attempts to measure variation between individuals is that it is often unclear to what extent sharing is in fact flexibility. *Flexibility in identification* refers to context-dependent variation in the relationship between terminal lexical items (or categories) and scientific taxa. By this I am not referring to situations in which an informant may label a taxon in a particular way because of

OMNISCIENCE AND IGNORANCE

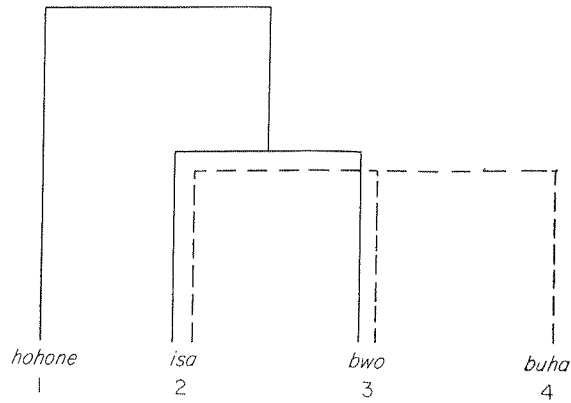


FIGURE 2. Alternative groupings for four Nuauulu reptile categories represented as a taxonomy.

inadequate stimuli, 'pseudo-compliance', uncorrected errors, or lack of specificity (see above), but to regular rule-guided variation according to context. For example, certain Nuauulu categories for skinks (*poso*) may be used interchangeably for the same scientific taxa, or several taxa may be described on different occasions by the same label (Ellen *et al.* 1977; cf. Hunn 1973: 45, 77; Hays 1974: 284). The term employed depends very much on the contrasts it is necessary to make in a particular speech situation. Logically, we should also admit at this point variation in identification governed by rules appropriate to definable social contexts—ritual, economic and so on—but I have no information to suggest that the Nuauulu make these discriminations. What is certain, though, is that classifications may vary according to rules appropriate to specific social contexts (Ellen 1975; cf. Rosaldo 1972). I discuss this matter further below.

CAUSES OF VARIATION IN KNOWLEDGE, IDENTIFICATION AND CLASSIFICATION

That informants vary in their empirical knowledge of animals, the identifications they make of particular species, and the broader categories to which they assign them is by now palpable. What is still unclear is the sociological significance to be attached to this in the analysis of a particular body of material. From the analysis of Nuauulu classification of amphibia and reptiles, it is clear that *in general* Nuauulu identifications are reasonably consistent, that the degree of sharing is extensive, and that the range of flexibility is confined, although apparently not so impressively as among the Kalam (Bulmer, Menzies & Parker 1975: 274). There is, then, a hard core of intersubjective agreement (cf. Harris 1964: 158). At least, this appears to be the case, given the known limitations of techniques and evidence.

In certain areas, however, Nuauulu identifications are remarkably inconsistent, and classifications are variable. This can be attributed, in part, to a relative absence of cognitive sharing between different subgroups, but there are also determinants affecting the group as a whole. For example, it has been suggested that migration, resettlement, and economic change have combined to alter the cultural significance of frogs, with the result that while the names remain, the experience necessary to employ them constantly, consistently, and perhaps also accurately is lacking (Ellen *et al.* 1976). There are also cases where the employment of 'loose labels' is intentional (see above). But what can be said in general about the conditions which affect variation?

Here I isolate four basic sets of variables, although none appear to be acting independently. The groups do not, therefore, indicate mutual exclusiveness. They are:

- (1) informant variability (according to age, gender, status, motivation, and so on);
- (2) contextual variability (linguistic, social, ecological and economic contexts);
- (3) methodological indeterminacy.

(1) *Informant variability*

(a) *Age*. Nuauulu children acquire knowledge of animal species gradually and unevenly as they obtain familiarity with an increasingly wider geographical area. There is an initial bias towards the domestic context: they become familiar with animals of and around the village and with game animals through meat brought back from the hunt. Some individuals have to wait until they are quite old before they see (let alone recognize) certain animals, and some young initiated men have never seen civets such as *Viverra tangalunga* and *Paradoxurus hermaphroditus*. The implication of this seems to be that because an individual has to acquire knowledge of animals, the older he or she is the larger the lexicons, the greater the experience, and therefore the more expert the identification (cf. Hays 1974: 209-19, 254). However, quite apart from the interference of extraneous variables associated with socialization - the development of skills, gender, status and so on - age ultimately brings with it physical and mental disability, if not in terms of finite knowledge and experience, then in ability to perceive, recall and connect.

(b) *Gender*. Differences in socialization between the sexes affect the relative knowledge of different kinds of fauna. The knowledge of men and women regarding different domains varies according to the social division of occupations. Women tend to have a more detailed knowledge of shellfish, freshwater fish and crustacea, and those of the littoral than men, owing to their pre-eminent role in fishing activities. Men, on the other hand, are more expert when it comes to

forest fauna. The marked influence of gender on plant knowledge, the number and proportion of names shared, has been confirmed empirically for the Ndumba of the New Guinea highlands by Hays (1974: 224, 254). With no permanent residential segregation of adult males and females among the Nuaulu, one might expect the pattern to be somewhat less marked. Whether or not this is the case has still to be determined.

(c) *Status and kinship position.* While knowledge of natural history may itself give an individual status, those persons who have it by virtue of ascription or some achievement (reputation in warfare and headhunting, in terms of personal wealth, as a shaman, magician, wise man, and so on) may be regarded as authorities in this sphere. This is so even if, in any objective sense, they are patently not. Bulmer & Tyler (1968: 347) have pointed out that a leading man cannot afford *not* to be an authority on any topic. With the Nuaulu, as with the Kalam, in the great majority of disputes over botanical or zoological identification, such persons are usually deferred to, at least in their presence, though younger men and boys occasionally continue to assert divergent opinions in their absence (Ellen 1975: 220). All this merely confirms what is already known of the role of political and ideological hegemony in other fields of knowledge. Persons of status may also show a greater aptitude and have a greater opportunity for acquiring such knowledge than others. The whole thing is a self-amplifying process of positive feedback, as with the relationship between education and status in stratified societies.

Status may also be relevant in another sense, in that for totemic and other clan-specific reasons an individual may be prevented from having access to particular groups of animals or environments. For example, the clan Sonawe-ainakahata is forbidden to have anything to do with the sea, while the fact that snakes are totemic beings for Somori restricts greatly the knowledge of this group's natural history among its members. On the other hand, the esoteric knowledge of certain animals may be passed on within certain descent groups and families and not within others. Differential socialization experience may therefore be important.

(d) *Linguistic competence.* Most adult Nuaulu are at least bilingual, speaking, in addition to their own tongue, either Ambonese Malay or the language of Sepa-Tamilau. Many, particularly adult males, but also some women (particularly those marrying into the group), speak a third language, and a few speak more than this. The languages most commonly represented here are those of Nuelitetu, Jalahatan, and Amahai, although those of Waraka and Manusela are also found (Stresemann 1927).

Multilingualism affects the kind of variation in indigenous classification systems, but, as the knowledge of a second language is seldom as detailed as the first, its effects seem to be restricted to higher-order groupings. That is, terminal

categories (unless these are also primary) tend to be affected less than those for groupings of the kind 'things that fly', 'big game', 'snakes', and so on.

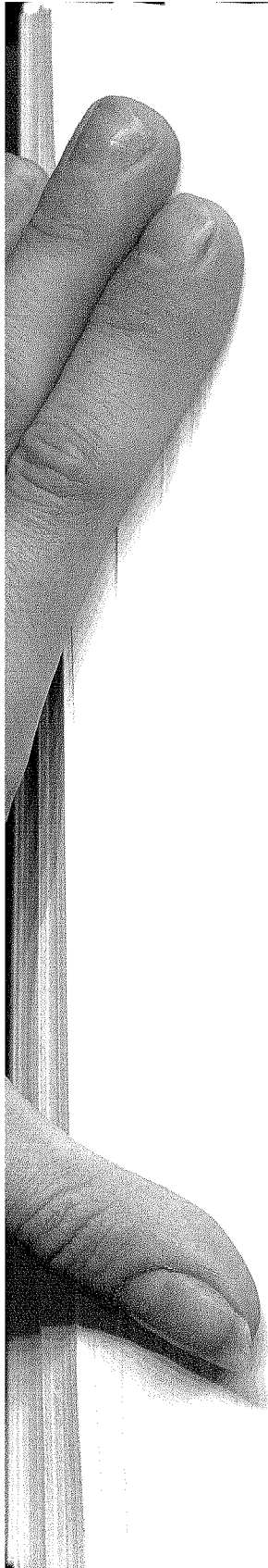
(e) *Ability and special skills.* Apart from the social determinants discussed above, knowledge of fauna may be affected by physical and technical ability. To a considerable degree physical ability is tied up with age, the learning of special sensory and motor skills, and their deterioration. Physical disabilities of other kinds may also be relevant, particularly those that impair travel and the senses. Pwete, a crippled polio victim in Ruhuwa village, had little knowledge of natural history outside the village, although this was compensated for by knowledge which her permanently sedentary existence allowed her to accumulate. Aharena, a deaf-mute in the same village, suffered from difficulties in communication, although she was sensitive to those aspects of natural history that her effective senses encouraged.

Mental ability must clearly also be relevant, though there was no way of judging its effect except by very crude means, and then its existence as an independent variable is extremely doubtful. It was clear, however, that certain individuals possessed skills which made them knowledgeable in particular areas of natural history. Generally, there were good social reasons for the distribution of these skills. The limitation of knowledge by status is one of these, marital state is another, and family tradition is yet another. But it is also the case that particular skills, however originally acquired, are often developed through interest and desire to concentrate on those areas where competence is high. There may also be other factors involved, such as contact with outsiders and other individuals possessing special skills. Opportunities for communication of information from other people are always important in this connection.

Two examples of this kind of situation should suffice. Naupati Matoke, an unmarried man, was particularly renowned as a hunter, but that he could devote so much of his time to this activity was largely due to his lack of kinship and affinal responsibilities. As a result, his knowledge of the natural history of the more-distant forest regions was probably unparalleled. A rather different case was Komisi Somori, the village headman of Ruhuwa, who spent a large proportion of his time engaged in inshore fishing. First, this was possible because there were no clan restrictions preventing him from doing so. Secondly, his status as headman required him to be constantly within easy reach of the village to entertain travellers and other guests. Hunting, forest collecting, and gardening activity would not be compatible with this. As a result, he had perhaps the most detailed and accurate knowledge of marine fauna of anyone in the village.

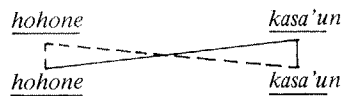
#### (2) *Contextual variability*

Different classifications, or identifications, may be elicited according to different contexts: linguistic, social or ecological (Ellen 1975). Put differently, a category



(at whatever level) may belong to several superordinate categories, and many criteria may be involved in grouping: morphological, anatomical, utilitarian, or ecological (cf. Friedberg 1971). Conflicting classifications may emerge, depending on whether the context is habitat, morphology, behaviour, or ideology.

(a) *Variation by linguistic and social context.* Little will be said here on variation according to linguistic context, since its ubiquity has already been the subject of much discussion (Whiteley 1966). Some examples from the Nuauulu corpus include the usage of uninomials and polynomials for the same taxon, depending on whether the identity of the animal is understood, and variable substitution of terms in a class inclusive relationship. Thus, in the first case, it is in order to speak of *nimasasai* (the gecko) when what is actually meant is *nimasasai numa* ('the house gecko', *Hemidactylus frenatus*). An example of the second case is that while in one elicitory context it may be appropriate to speak of *kasa'un* (*Calotes cristatellus*) as a type of *hohone*, in another it may be equally in order to talk of *hohone* (*Draco lineatus*, the 'flying lizard') as a type of *kasa'un*:



Such variable practices appear to derive from the priority or frequency of particular terms in a stretch of discourse. If a conversation opens with a discussion of (or reference to) *hohone*, then *kasa'un* is a type of *hohone*; and vice versa. This usage is also reported for *isa* (*Hydrosaurus amboinensis*) and *bwo* (*Varanus indicus*, the water monitor); *nopa hanaie* (*Eugongylus rufescens*) and *nopa inai* (*Tiliqua gigas*); and *poro-poro* (tree frogs) and *notu* (ground frogs).

I have presented two extended examples of classificatory flexibility according to social context elsewhere (Ellen 1975), but variation according to ritual, general conversation, and economic activities are important generally. Variation according to differing ecological and economic contexts is considered below. Here I shall confine my remarks to one context where classifications associated with it often tend to be at considerable variance with those derived from other principles.

Classification by mythic origin, in particular, may conflict strikingly with those apparently based on a polythetic combination of morphological, behavioural, and ecological features, although, to some extent, mythic knowledge may be absorbed into such classifications. For example, mythically, many lizards are ontologically close to snakes. Thus, most skinks (and possibly also dragon lizards, geckos, and monitors) are said to be derived from the reticulated python. Also, the term *nanate* (death adder) may also be used more broadly to include the skinks *Tiliqua gigas* and *Eugongylus rufescens*, grouping together reptiles for which there is a marked aversion. However, in terms of other criteria, snakes are quite separate from lizards. They are distinguished by the shape of the head, the tongue,

TABLE 2. Results of a test in which 25 informants were asked to sort picture cards into 'natural' groups based on degree of similarity

	notu, ground frogs	poro-poro, tree frogs	peku, Amboinan box terrapin	enu ikai, leatherback turtle	buha, crocodile	nimasasai, geckos	kasa'un, Calotes crystalinus	isa, Hydrosaurus amboinensis	beco, monitor lizard	poso, skinks	tekene, snakes
1. notu	—										
2. poro-poro	25										
3. peku	2	2									
4. enu ikai	1	2	22								
5. buha	1	1	11	9							
6. nimasasai	—	—	1	3	8						
7. kasa'un	1	1	1	2	1	8					
8. isa	1	1	2	4	15	17	5				
9. beco	1	1	2	4	12	20	5	23			
10. poso	1	1	1	2	9	21	5	15	18		
11. tekene	—	—	—	—	—	3	1	—	—	3	—

taxa to those for other animals. These matters must be deferred to some future occasion. The aim here is simply to suggest a means of representing statistical variation in the views of informants on the similarities and differences between categories.

3) *Variability in the properties of different kinds of natural entity*

Informants appear to be better at identifying some groups than others, whether these are defined phylogenetically, ecologically, or culturally. We have already seen that this may be related to the opportunity and ability of various individuals and groups within Nuaulu society. Thus, women have a greater knowledge of freshwater fish because fishing and collecting along inland waterways is pre-eminently female work. Similarly, adult men have a far better knowledge of nature forest fauna than women, due to the opportunities provided by hunting and trapping. However, some variation must be attributed to the objective characteristics of the animals themselves, rather than to their utilitarian values or variation between types of informant. Animals with a distinctive behavior or morphology (and perhaps also greater size) will tend to be discriminated positively and elaborated more in classifications than those that are undistinctive in these respects. Population size and any harmful effects they may have on humans and on natural species of value to humans may also be important.<sup>3</sup> For example, Nuaulu informants, as a whole, were much more consistent and able in identifying molluscs than some other animals. The reasons for this can be attributed to the facts that:

- (a) diagnostic characteristics are almost entirely morphological and do not change in the live or dead states or in different contexts, including cooking;
- (b) they are easily preserved over long periods; and
- (c) there are a lot of them around (in the rivers, in the house, on middens), so that informants have opportunities to make more frequent identifications than for most other animal groups.

So, it is possible that the Nuaulu are so much better at identifying reptiles than frogs because there is a greater *morphological* diversity among the former. Their apparent failure as amphibian biologists stems in part from social, economic, and ecological factors (see above) but also from the fact that specific and even generic) differences are more finely distinguished phylogenetically – subtle variation in colour, distinctiveness of tympanum, behaviour, and so on. Many of the creatures are even difficult for professional herpetologists to identify

---

<sup>3</sup> In general, if insignificant animals are at all in significant relationships with other animals which are significant in themselves, then they will tend to be ordered in classificatory systems (Bulmer 1975: 12–13).



at this level, if superficial morphological characteristics alone are taken into account. In view of this, it is not surprising that the Nuaulu have difficulty.

(4) *Genuine variation or methodological indeterminacy?*

One difficulty which very rapidly emerges in empirical research on classifications is the extent to which conflicting information is to be interpreted as significant variation rather than carelessness or oversight on the part of the ethnographer or weaknesses in the techniques employed. The importance of careful work, therefore, cannot be over-stressed. Because he or she is an outsider, and because his or her knowledge of the language is inevitably limited, the ethnographer may tend to simplify, to formalize or play down variation where this is unwarranted. On the other hand, it may be given undue emphasis. No doubt these things are always a source of error (Ellen 1975: 218-9; Bulmer 1969), but how do we go about ascertaining its extent (Bricker 1974: 69-70)?

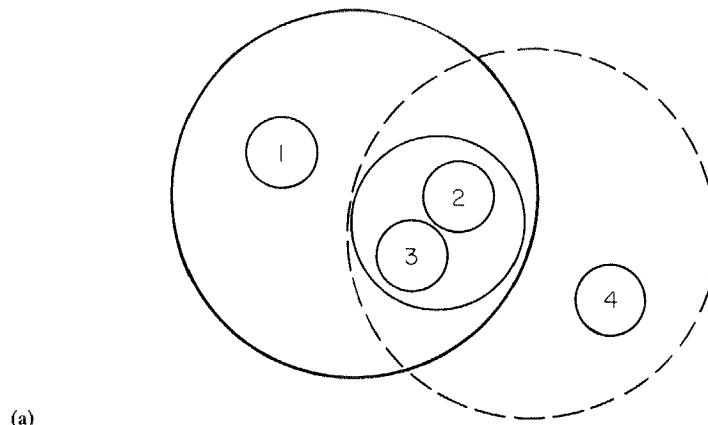
Much error of this kind stems from the difficulty of determining truly *individual* variation under fieldwork conditions. On the one hand, it is generally impossible to achieve experimental standards comparable to those achieved in psychological testing carried out under laboratory conditions. On the other, it might be argued that attempts to emulate such procedures introduce yet other artifacts which distort results. The partial knowledge on which ethnobiological inventories are based is often that of small groups rather than individuals. It is important to question the very value of looking at *individual* variation, when, in the real world, many decisions regarding identification and classification are arrived at collectively. The collectivity is not 'the society' but much smaller groups of individuals, the composition of which is constantly changing (but where decisions are affected by the range of determinants already discussed). Answers given by any one informant in a session may be influenced by the comments (or even the mere presence) of others. In fact, he or she may actively solicit the comments of others, and this may be accepted as the normal way of going about things. Which of the conflicting identifications given for a single specimen is accepted as definitive depends on the group of persons assembled. What is regarded as an 'error' is determined by what the ethnographer decides is the significant variation of interest, for even 'errors' can have rational explanations. They are also the inevitable logical result of attempts to construct composite 'folk models' of culture (Hays 1974: 131).

Individuals themselves are rarely consistent in the statements they make in elicitory contexts. They may initially provide one name, or indeed several, reject them, and finally settle on something quite different. This is the identification which must normally be accepted as 'definitive', although it is possible that further observation, consideration, or consultation may lead the informant to change his or her mind again. The problem with work of this kind is that it

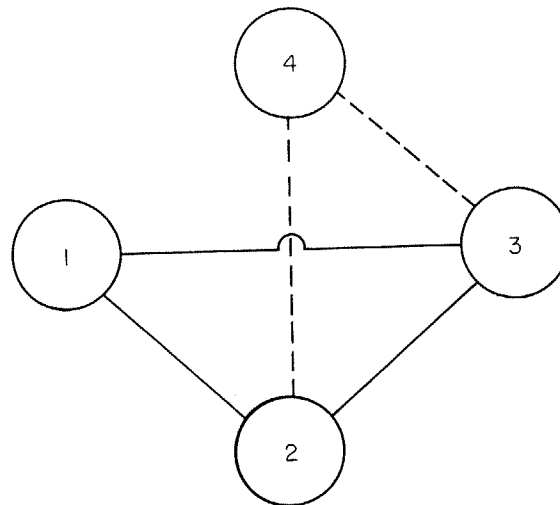
OMNISCIENCE AND IGNORANCE

requires definitive answers for analytical reasons, while it is known that the processes generating the answers are capricious.

Variability may also be an artifact of taxonomic method (Friedberg 1968, 1970; Ellen 1979), or other modes of representation. If it is assumed that informants operate according to a rigorous logic of hierarchy and class inclusion in which transitivity is encoded, then 'variation' is inevitable. Thus, although I have chosen to represent the relationships between Nuaulu categories for amphibians and



(a)



(b)

FIGURE 3. Alternative groupings of four Nuaulu reptile categories represented (a) as a Venn diagram, and (b) as a network. (Numerals refer to categories in Fig. 2.)

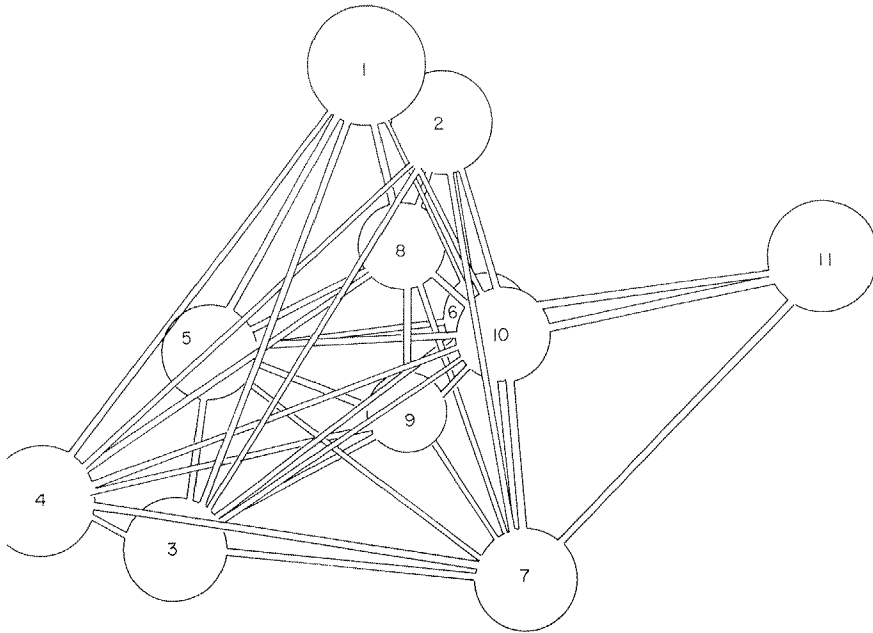


FIGURE 4. Distance matrix for selected Nuaulu reptile and amphibian categories, represented as a three-dimensional physical model. (Categories are shown as spheres connected by rods the length of which indicates relative cognitive proximity. Numerals refer to categories listed in Table 2.)

reptiles through a series of taxonomic charts (for reasons of convenience and communicative effect), this has given rise to some representational difficulties. Relationships which might have been more simply and accurately represented in other ways have had to be illustrated as interlocking hierarchies suggesting deviant practice. An example of this is the relationship between the categories *buha* (crocodile), *hohone* (flying lizard), *bwo* (monitor), and *isa* (*Hydrosaurus amboinensis*). As part of a simplified taxonomic chart, this might be represented as shown in Fig. 2. Where taxonomies can be shown to be part of the classificatory equipment used by informants, then this must be evident from representations, but their use across the board reifies certain connections and variations while concealing others. One possible alternative, already mentioned, is the use of Venn diagrams. The relations portrayed in Fig. 2 would then look something like Fig. 3a. Another possibility is simply to represent the relationships as a network as in Fig. 3b.

Fig. 4 goes one step further by depicting the relationships three-dimensionally, although in this case it relates to all the categories included in Table 2. Those spheres that are closest indicate pairs of categories whose relationship is most

agreed upon. There is no reason to assume that classificatory space is two-dimensional, and good reason to doubt it. Sometimes even three dimensions are insufficient to express complex relations between categories. All too frequently, elicited data is distorted through attempts to represent complex interlocking relationships in elegant diagrams which have height and breadth, but no depth. I have made some effort to discover suitable graphical means of displaying the spatial relationships between categories but the task is by no means as simple as it may seem. The present technique (derived from Sokal & Sneath 1963: 203) has some didactic value; it at least gives some impression of cognitive distance and connectedness. It is still far from satisfactory, as perhaps all such attempts are bound to be. They compress and express data in ways which bear only a distant relation to actual classificatory and social practice.

#### CONCLUSION

There is considerable variation among the Nuauulu in the range and accuracy of individual and sub-group knowledge of animals, their names, and how these should be assigned to categories. Individuals vary in the experience of these groups due to social differentiation, context, and, to some extent, because of the characteristics of the animals themselves. It is personal experience and the exchange of personal experiences that are important at both the level of identification and in much assignation of animals to more inclusive categories. The authority of tradition, socialization, and the social distribution of power and of formal status may certainly affect allocations and arrangements, particularly at the more abstract level. In some cases, such forces may guarantee a single favoured pattern, but, as a general rule, the framework has to be flexible enough to allow individuals to fit into it as much as possible of their unique experience (Bulmer 1975: 11).

Variability is what we must expect, diversity is part of the system; homogeneity is not only not necessary, or even possible (Hays 1974: 204, 305), but perhaps is also not always desirable. This is not only because people's knowledge, interests, and abilities in identification and classification vary, but also because people constantly use categories, without formally defining them, in ways which suggest different, even opposite, properties on different occasions (Bourdieu 1977: 110). It is for this reason that I am as wary of the theoretical pre-eminence of the 'individual' as I am of the pre-eminence of 'culture'. It is, of course, individuals (not cultures) who perceive (Hays 1974: 99), but they do so in a world where social interaction is continuous and determinate. The lexicon, the knowledge, and classificatory construct of the individual can all too easily be reified through the careless use of the language of description and graphic representation. 'Mental maps', complex semantic models encased in the minds of informants, are just as improbable at the level of classification as are static and

pictorial mental images at the level of identification. Undoubtedly, people do carry around constructs of some kind, but, more importantly, they carry the ability to constantly recall, dissect, and combine them according to changing exterior social and environmental stimuli and interior reflection. It is necessary to represent classifications and their variation in particular ways, for analytical and didactic purposes, but we only do so at the expense of understanding the full richness of the relations between classifying as a process, social practices in which these take place, and the classifications which are their products.

## REFERENCES

- Berlin, B. (1972). Speculations on the growth of ethnobotanical nomenclature. *LinS* 1 (1). 51-86.
- Berlin, B., Breedlove, D. & Raven, P. (1974). *Principles of Tzeltal plant classification: An introduction to the botanical ethnography of a Mayan-speaking people of the highland Chiapas*. New York: Academic Press.
- Bloch, M. (1971). The moral and tactical meaning of kinship terms. *Man* 6 (1). 79-87.
- Bourdieu, P. (1977). *Outline of a theory of practice* (trans. Richard Nice). (Cambridge Studies in Social Anthropology No. 16.) Cambridge: Cambridge University Press.
- Bricker, V. R. (1974). Some cognitive implications of informant variability in Zinacanteco speech classification. *LinS* 3. 69-82.
- Bulmer, R. N. H. (1969). Field methods in ethno-zoology with special reference to the New Guinea highlands. University of Papua and New Guinea: mimeographed.
- (1970). Which came first, the chicken or the egg-head? In J. Pouillon & P. Maranda (eds), *Échanges et communications: Mélanges offerts à Claude Lévi-Strauss*, Vol. 2. The Hague-Paris: Mouton.
- (1974). Memoirs of a small game hunter: On the track of unknown animal categories in New Guinea. *Journal d'Agriculture Tropique Botanique Appliquée* 21 (4-6). 79-99.
- (1975). Folk-biology in the New Guinea highlands. *Social Science Information* 13 (4/5). 9-28.
- Bulmer, R. N. H. & Menzies, J. I. (1972-3). Karam classification of marsupials and rodents. *Journal of the Polynesian Society* 81 (4). 472-99; 82 (1). 86-107.
- Bulmer, R. N. H., Menzies, J. I. & Parker, F. (1975). Kalam classification of reptiles and fishes. *Journal of the Polynesian Society* 84 (3). 267-308.
- Bulmer, R. N. H. & Tyler, M. J. (1968). Karam classification of frogs. *Journal of the Polynesian Society* 77 (4). 333-385.
- Coxon, A. (1977). Classification and hierarchy construction as methods of eliciting subjective information. Paper presented at the Social Science Research Council seminar on 'Classification systems and the anthropology of knowledge' at the University of Kent at Canterbury, June 1977.
- Ellen, R. F. (1975). Variable constructs in Nuauulu zoological classification. *Social Science Information* 14 (3/4). 201-28.
- (1976). Restricted faunas and ethnozoological inventories in Wallacea. *Working Papers in Southeast Asian Studies*. Second series, No. 10. School of Oriental and African Studies, University of London.
- (1979). Formal categories in their social context. In R. F. Ellen & D. Reason (eds) *Classifications in their social context*. London: Academic Press.
- Ellen, R. F., Stimson, A. F. & Menzies, J. I. (1976). Structure and inconsistency in Nuauulu categories for amphibians. *Journal d'Agriculture Tropique Botanique Appliquée* 23 (7-12). 125-38.
- (1977). The content of categories and experience: The case for some Nuauulu reptiles. *Journal d'Agriculture Tropique Botanique Appliquée* 24 (1). 3-22.

- (1977a). On some further aspects of Nuauulu classification of snakes and other reptiles. Unpublished manuscript.
- Friedberg, C. (1968). Les méthodes d'enquête en ethnobotanique. *Journal d'Agriculture Tropic Botanique Appliquée* 15 (7-8), 297-324.
- (1970). Analyse de quelques groupements de végétaux comme introduction à l'étude de la classification botanique Bunaq. In J. Pouillon & P. Maranda (eds), *Échanges et communications: Mélanges offerts à Claude Lévi-Strauss*, Vol. 2. The Hague-Paris: Mouton.
- (1971). Aperçu sur la classification botanique Bunaq (Timor central). *Bulletin Botanique Française* 118, 255-62.
- (1974). Les processus classificatoires appliqués aux objets naturels et leur mise en évidence: Quelques principes méthodologiques. *Journal d'Agriculture Tropic Botanique Appliquée* 21 (10-12), 313-34.
- Gal, S. (1973). Inter-informant variability in an ethno-zoological taxonomy. *Anthropological Linguistics* 15, 203-19.
- Gardner, P. M. (1976). Birds, words and a requiem for the omniscient informant. *American Ethnologist* 8, 446-68.
- Hage, P. & Hawkes, K. (1975). Binumarien color categories. *Ethnology* 14 (3), 287-300.
- Harris, M. (1964). *The nature of cultural things*. New York: Random House.
- (1968). *The rise of anthropological theory*. New York: Crowell.
- (1970). Referential ambiguity in the calculus of Brazilian racial identity. *Southwestern Journal of Anthropology* 26 (1), 1-14.
- Hays, T. E. (1974). Mauna: Explorations in Ndumba ethnobotany. Ph.D dissertation. Seattle: University of Washington.
- (1976). An empirical method for the identification of covert categories in ethnobiology. *American Ethnologist* 8, 489-507.
- Heider, E. R. (1972). Probabilities, sampling and ethnographic methods: The case of Dani colour names. *Man* 7 (3), 448-66.
- Henderson, J. & Harrington, J. P. (1941). *Ethnozology of the Tewa Indians*. (Bulletin of the Bureau of American Ethnology, No. 56.) Washington: Government Printing Office.
- Hunn, E. (1973). Tzeltal folk zoology: The classification of discontinuities in nature: Ph.D. dissertation. Berkeley: University of California.
- (1975). The Tenejapa Tzeltal version of the animal kingdom. *Anthropological Quarterly* 48, 14-30.
- (1975a). *Cognitive processes in folk ornithology: The identification of gulls*. Working paper 42. Language Behavior Research Laboratory, University of California, Berkeley.
- Malkin, B. (1958). Cora ethnozology, herpetological knowledge: A bio-ecological and cross-cultural approach. *Anthropological Quarterly* 31 (3), 73-90.
- Manning, P. K. & Fabrega, H., Jr. (1976). Fieldwork and the 'new ethnography'. *Man* 11 (1), 39-52.
- Needham, R. (1975). Polythetic classification: Convergence and consequences. *Man* 10 (3), 349-639.
- Rosaldo, M. Z. (1972). Metaphors and folk-classification. *Southwestern Journal of Anthropology* 28 (1), 83-99.
- Sankoff, G. (1971). Quantitative analysis of sharing and variability in a cognitive model. *Ethnology* 10, 389-408.
- Simpson, G. G. (1961). *Principles of animal taxonomy*. New York: Columbia University Press.
- Sokal, R. R. & Sneath, P. H. (1963). *Principles of numerical taxonomy*. San Francisco-London: W. H. Freeman.
- Sperber, D. (1975). Pourquoi les animaux parfaits, les hybrides et les monstres sont-ils bon à penser symboliquement? *L'Homme* 15 (2), 5-34.
- Stresemann, E. (1927). Die Lauterscheinungen in den amboinischen Sprachen. *Zeitschrift für Eingeboren-Sprachen* 10.
- Tyler, S. A. (ed.) (1969). *Cognitive anthropology*. New York: Holt, Rinehart & Winston.
- Wallace, A. F. C. (1970). *Culture and personality* (2nd ed). New York: Random House.
- Whiteley, W. H. (1966). Social anthropology, meaning and linguistics. *Man* 1 (2), 139-57

APPENDIX: SIMPLE MEASURES OF VARIATION  
IN LEXICAL FIELDS

(David Reason, *University of Kent at Canterbury*)

This appendix suggests a number of measures which may be readily calculated in respect of data on lexical usage to provide an indication of the degree of variation in such usage. It introduces and briefly discusses five straightforward measures of variation or consensus which apply to the case where any informant has been constrained to give only one response per stimulus item. None of these measures is, strictly speaking, a statistical measure, although  $p$  is almost so in terms of its logic and the  $D$ -measures have a form familiar to that of a wide class of elementary statistics. Each measure in fact generates a numerical index which relates to an arithmetical property of the data in question: the measures are not in general estimates of population parameters. The use is heuristic, exploratory and comparative: by systematically reducing the information which comprises the data, the measures should help the investigator to become more aware of the salient characteristics of the data to hand.

Throughout, we assume  $N$  informants ( $I_1, I_2, \dots, I_N$ ), and a total of  $L$  lexical items ( $l_1, l_2, \dots, l_L$ ), each item  $l_i$  being 'supported' by  $n_i$  informants.

*The 'simple majority' measure, M*

In computational terms, the simplest measure of agreement in lexical choice of use in interspecific comparisons of lexical agreement is *the proportion of informants opting for the most frequently cited item*,

$$M = \frac{\max(n_i)}{N} \quad (0 < M \leq 1).$$

$M$  enables one to make interspecific comparisons since it is standardized to a uniform number of informants; it also allows interspecific comparisons of order of 'popularity' for ranges of lexical items if  $M$  is computed for all items and not just those most frequently cited.

$M$  has all the drawbacks as a general measure of agreement as has a simple-majority-rule voting system as a device for representing 'the general will': in particular,  $M$  fails to make use of information concerning the number of lexical items available and the distribution of support for these less popular, but 'competing', items. This is particularly evident in considering the problem of interpreting the converse measure,  $(1-M)$ , as a measure of dissensus or disagreement, when it becomes crucial to know whether the opposition is united or divided (and to what degree) – and this information is excluded from the outset. Moreover, because  $M$  fails to incorporate information on the distribution of support, it is a relatively coarse measure, equating patterns which would

usi  
(m  
the

M

If,  
N  
(N  
fre  
res  
di

H  
ch  
in  
oc  
in

H  
uj  
de  
di

ol  
is

(i  
b

It  
at

w

A  
V

a  
o

w  
l

o

usually be better kept separate. In short,  $M$  is essentially an item measure (making only the very faintest of acknowledgements to the distribution in which the chosen item is found), not a structural measure.

*Measure of enhancement in prediction,  $\Delta p$*

If, having gained information concerning the identification of a specimen from  $N$  informants, we were to predict the identification made by a hypothetical  $(N+1)$ st informant, we would be justified in predicting the hitherto most frequently cited item: provided our group of informants was (in all material respects) representative of the culture as a whole, the probability of our prediction being correct would be close to

$$\frac{\max(n_i)}{N} = M.$$

Had we simply known how many lexical items the  $(N+1)$ st person could have chosen from, and been ignorant of the distribution of responses given by our informants, we should have assigned each lexical item the probability  $(1/L)$  of occurring next time. Our knowledge of the distribution of responses has thus increased our accuracy of prediction by

$$\Delta p = M - \frac{1}{L} \quad (0 \leq \Delta p < 1).$$

Hence  $\Delta p$  reduces the simple majority measure  $M$  by an amount which depends upon the number of lexical items, and as the number of items increases, so  $1/L$  decreases and  $\Delta p$  tends towards the value of  $M$ .  $\Delta p$  does not depend upon the distribution of responses in the fullest sense, however, since it takes no account of the relative strength of support for competing items. Moreover, its validity is strictly dependent upon the group of informants forming a representative (in the statistical sense) sample of the population at large, and so it should only be used where there is a sufficient number of informants (certainly not less than 10, but the precise number depending upon the number of lexical items accumulated). Like  $M$ ,  $\Delta p$  may be regarded as a measure of the degree of support which the most frequently cited items commands.

*Measure of qualitative variation,  $V$*

We may suppose that a lack of agreement as to the correct lexical item for a given species may form the basis for argument or negotiation. The measure of qualitative variation compares the actual extent of such potential disagreement with the maximum conceivable extent (Mueller, Schuessler & Costner 1970; Nachmias & Rosenbloom 1973).

For any two lexical items, one of which is endorsed by  $n_i$  informants and the other by  $n_j$ , the number of distinct *pairs* of individuals who would disagree over



'what the thing shall be called' is  $n_i n_j$ : hence the total number of such disagreements, considering all the lexical items, is

$$\sum_{i < j} n_i n_j.$$

The maximum number of disagreements would occur if each lexical item received the same level of support, i.e.  $(N/L)$ , and is therefore

$$\sum_{i < j} \left(\frac{N}{L}\right)^2 = L \left(\frac{L-1}{2}\right) \left(\frac{N}{L}\right)^2,$$

since there are  $L(L-1)/2$  distinct pairs to be found in  $L$  items. Hence,

$$V = \frac{\sum_{i < j} n_i n_j}{\frac{L(L-1)}{2} \left(\frac{N}{L}\right)^2} \quad (0 \leq V \leq 1)$$

$V$  is a structural measure, taking into account all features of the distribution of responses:  $V = 0$  would indicate no variation, *but* in the application considered in this appendix,  $V = 0$  is strictly unobtainable since  $V$  is undefined for a unanimous response - we therefore *assign*  $V = 0$  to the 'nem. con.' situation.  $V = 1$  indicates that all items receive equal support.

$V$  is predicated upon the legitimacy of taking pairwise encounters alone as the locus of disagreement, and its interpretability is dependent upon the rationale for such a stipulation in any particular empirical investigation.

#### *Measures of deviation from a criterion distribution, $\bar{D}$ and $\hat{D}$*

All measures rely for their interpretation on some conceivable state of affairs which serves as a standard and against which an empirically obtained distribution is compared. In the case of  $M$ , the standard is that of unanimity; of  $\Delta p$ , the 'random' allocation of individuals to lexical items; whilst  $V$  refers to a property of a homogenous distribution of individuals to lexical items. In each of these cases the measure is a linear function of the specified contributing conditions, in that each pairwise encounter 'counts' just as much as does any other, and so on.

With the  $D$ -measures, the procedure is: a criterion frequency distribution is specified, and the sum of the squared deviations of the empirically obtained distribution from the criterion distribution forms the basis of the measure. Squaring the deviations weights large deviations relatively more than small deviations: for example, the relative deviations of 2 and 3 give squared deviations of 4 and 9 respectively, the relative difference of the latter being considerably

reater than in the unsquared case. So, if  $f_i$  is the criterion frequency for lexical item  $l_i$ , we have a class of  $D$ -measures:

$$D = \frac{\Sigma(n_i - f_i)^2}{\Sigma(n_i^* - f_i)^2},$$

where the  $\{n_i^*\}$  are chosen so as to maximize the denominator. The two most useful criterion frequencies are the flat distribution

$$\{\bar{f}_i\} = \left\{ \frac{N}{L}, \frac{N}{L}, \dots, \frac{N}{L} \right\}$$

and the peaked distribution:

$$\{\hat{f}_i\} = \{N, 0, 0, \dots, 0\}.$$

The flat distribution amounts to a hypothesis of no cultural preference for any particular lexical item, as if informants were being asked for arbitrary choices in a situation of free variation. The peaked distribution would obtain if only one lexical item (in practice, that item with the greatest empirical frequency) were current among the informants. The relevant  $D$ -measures, giving the degree of deviation from the flat and peaked criterion distribution, are then:

$$\begin{aligned} \bar{D} &= \frac{\Sigma(n_i - \bar{f}_i)^2}{\Sigma(\bar{n}_i^* - \bar{f}_i)^2} = \frac{\Sigma\left(n_i - \frac{N}{L}\right)^2}{\Sigma\left(\bar{n}_i^* - \frac{N}{L}\right)^2} \\ \hat{D} &= \frac{\Sigma(n_i - \hat{f}_i)^2}{\Sigma(\hat{n}_i^* - \hat{f}_i)^2} = \frac{(n_{\max} - N) + \sum_{j \neq i} (n_j)^2}{\Sigma(\hat{n}_i^* - \hat{f}_i)^2} \\ &= \frac{\Sigma n_i^2 + N(N - 2n_{\max})}{\Sigma(\hat{n}_i^* - \hat{f}_i)^2}, \end{aligned}$$

where  $n_{\max}$  is the greatest frequency with which an item is elicited, i.e.  $\max(n_i)$ .

Now, if we specify the flat distribution, the largest sum of squared differences is obtained if  $\{\bar{n}^*\}$  follows the peaked distribution: and, necessarily, vice versa. The denominators of  $\bar{D}$  and  $\hat{D}$  are thus identical: in fact,  $N^2(1 - 1/L)$ .

Hence the measure of deviation from a flat distribution,

$$\bar{D} = \frac{\Sigma\left(n_i - \frac{N}{L}\right)^2}{N^2(1 - 1/L)}$$

the measure of deviation from a peaked distribution,

ROY F. ELLEN

$$\hat{D} = \frac{N(N - 2n_{\max}) + \sum n_i^2}{N^2(1 - 1/L)}$$

$\bar{D}$  is computationally undefined when the empirical distribution has only one item: we introduce the convention that  $\bar{D} = 0$  in such a case. In general,  $\bar{D}$  tends to zero the closer is the empirical distribution to a flat distribution, and tends to one the more peaked is the empirical distribution.  $\hat{D}$  is interpreted in like fashion, and for similar reasons we introduce the convention that, for a single item,  $\hat{D} = 0$ . These conventions are, of course, pure mathematical conveniences and have no substantive importance: where only a single item is reported, any attempt to calculate an index of variation is quite simply absurd.

Interestingly, it is not  $\bar{D}$  which is the natural converse of  $\hat{D}$ , but  $V$ . It is absolutely straightforward to show that  $D = 1 - V$ ; and, indeed, this is as it should be, since  $D$  tends to zero the closer the empirical distribution is to the flat distribution, in which case  $V$  is tending to 1, since this is precisely the distribution which maximizes pairwise differences.

The behaviour of these 'five' measures (they produce only four independent values) is illustrated by Table 3, where they are calculated from fictitious data.

TABLE 3.

Species	$l_1$	$l_2$	$l_3$	$l_4$	$N$	$L$	$M$	$\Delta p$	$V$	$\bar{D}$	$\hat{D}$
A	8	2	—	—	10	2	0.80	0.30	0.64	0.36	0.16
B	7	2	1	—	10	3	0.70	0.37	0.69	0.31	0.21
C	6	2	1	1	10	4	0.60	0.35	0.77	0.23	0.29
D	8	1	1	—	10	3	0.80	0.47	0.51	0.49	0.90
E	4	2	2	2	10	4	0.40	0.15	0.96	0.04	0.64
F	4	3	2	1	10	4	0.40	0.15	0.93	0.07	0.69

REFERENCES

- Mueller, J. H., Schuessler, K. F. & Costner, H. L. (1970). *Statistical reasoning in sociology*. Boston: Houghton Mifflin.  
 Nachmias, D. & Rosenbloom, D. H. (1973). Measuring bureaucratic representation and integration. *Public Administration Review* 33. 590-7.