Biodiversity and Conservation 14:2971–3001
DOI 10.1007/s10531-004-0256-4

Application of diversity indices to appraise plant availability in the traditional medicinal markets of Johannesburg, South Africa

VIVIENNE L. WILLIAMS*, EDWARD T.F. WITKOWSKI and Kevin Balkwill
*School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, PO Wits, 2050, South Africa; *Author for correspondence (e-mail: vivwill@planetac.co.za; fax: +27-11-403-1429)*

Received 4 June 2003; accepted in revised form 19 April 2004

Key words: Abundance, Diversity, Evenness, Quantitative ethnobotany, Species accumulation curve, Species richness, Traditional medicinal trade

Abstract. The lack of scientific rigour in analysing ethnobotanical surveys has prompted researchers to investigate ways of quantitatively describing their data, including the use of ecological diversity indices. There are numerous indices and measures available to describe sample diversity. Twenty-two measures of species richness, diversity and evenness were reviewed using six sets of ethnomedicinal data derived from 50 formal *muti* shop traders (of different ethnicities) and 100 informal street traders of traditional medicine in Johannesburg, South Africa, and a seventh data set from traders on the western boundary of the Kruger National Park, South Africa. The diversity measures were coupled with species accumulation curves to construct cumulative diversity curves used to determine the minimum viable sample size on which a diversity index should be based, and to better understand the differences in the relative diversities of the samples. Distinct differences in the relative abundance and diversity of plants sold by street traders and shop traders were evident. Species diversity and evenness was found to be higher in shops, thus resulting in a lower dominance in the sale of certain plant species compared to the street traders. A survey of an informal market should include no less than 35 research participants compared to no less than 20 for the *muti* shops. The use of selected indices of species richness (Margalef's), diversity (Shannon, Simpson's, Fisher's alpha, Hill's numbers) and evenness are recommended as a means of describing patterns exhibited within ethnobotanical data.

Introduction

An emergent trend in ethnobotanical studies has been the use of quantitative methods and models to describe patterns of plant use and availability in surveys or assessments of natural resources (e.g., Prance et al. 1987; Phillips and Gentry 1993a, b; Johns et al. 1994; Begossi 1996; Höft et al. 1999; Hanazaki et al. 2000; Luoga et al. 2000; Cunningham 2001; Wong et al. 2001), thereby allowing for a more rigorous statistical approach to the discipline. While a quantitative approach to the discipline is not always possible or even necessary, the benefits include: a greater depth to the understanding of the subject under investigation; a conscious attempt at reporting and refining methods employed to collect and evaluate the data; attention to sampling effort; economy of description and examination of patterns in the data; hypothesis testing; and the ability to question and describe more precisely the results of surveys.

Begossi (1996) first demonstrated how diversity indices were useful quantitative tools that could assist researchers analysing ethnobotanical data by allowing comparisons of diversity among different communities in different or similar environments. In ecology there are numerous indices available for exploring species diversity between different communities. Begossi (1996) demonstrated how the Shannon–Wiener index of diversity and evenness, and the rarefaction curve, might be used to compare differences in diversity, uniformity of species use and sampling effort in eight samples from South America, Thailand and Tonga. This paper broadens the spectrum of diversity indices to include 22 measures of richness, evenness and diversity. The goals are to: evaluate the performance of the indices in relation to samples of different sizes, and trader profiles and to examine the kind of information they provide; make recommendations on measures appropriate for quantifying ethnobotanical data; assess whether the survey sites were adequately sampled, and determine the minimum viable sample size on which a diversity measure should be based for the type of survey data collected; and lastly, to compare the species diversity of sites within the formal and informal sector, and thereby appraise plant availability within the ethnobotanical trade in the region.

Species diversity

Diversity indices are used to characterise species abundance relationships in communities (Ludwig and Reynolds 1988). The literature reveals a multifarious array of ecological indices, usually expressing species diversity as a single number (Magurran 1988). How can the appropriate measure be determined? The answer largely depends on the question the index is being used to answer, the component of diversity being measured, and whether the index is simple to use and understand.

Diversity measures take into account two factors: species richness (i.e., the number of species, *S*) and evenness/equitability (i.e., how uniformly abundant species are in a sample) (Magurran 1988). An 'index' of species diversity (also called an index of heterogeneity) incorporates both richness and evenness into a single value. Species diversity measures are broadly divided into four main categories: indices of species richness; indices of evenness; indices of species diversity/heterogeneity; and species abundance/distribution models (Ludwig and Reynolds 1988; Magurran 1988). These concepts may be translated into ethnobotanical terms to answer the following questions: (1) what is the species richness of plants used/sold within a sampled market or group of resource users?; (2) how does the species diversity of plants sold differ between different groups of traders?; (3) are the same plants sold by most traders?; and (4) is sampling effort adequate, i.e., have sufficient numbers of research/survey participants been interviewed?

Recent research by Colwell and Coddington (1995) and Gotelli and Colwell (2001) recommend the measurement and comparison of species richness by the use of taxon sampling or accumulation curves. The curves may be computed from *EstimateS* (Colwell 2001) software that computes randomised species accumulation curves and also several indices of diversity and the parameters to calculate others. Plotting the performance of the indices on a diversity curve demonstrates the performance of the indices and differences in relative abundance as sample size increases.

Study area

Johannesburg is located within a region of Gauteng Province called the 'Witwatersrand', the name given to an extensively urbanised axis of approximately 100 $\times$ 40 km and part of a geological super-group consisting of gold-bearing conglomerates (Lowrey and Wright 1987). The Witwatersrand emerged from a small mining town in the 1880s and labour for the mines were provided by mainly rural people in the migrant labour system. The ensuing rural–urban oscillation of Black labour from around the country enhanced the introduction of activities related to Black 'rural' culture (Dauskardt 1990, 1991). Traditional herbalism was incorporated into the developing urban mine culture to meet the needs of both the Black migrant labourers and the rapidly expanding, permanent urban population for traditional medicine (Dauskardt 1991). Various historical processes shaped the preponderance of different ethnic and language groups within sectors of the emerging South African capitalist economy and the traditional medicine trade. The Witwatersrand is South Africa's second largest market for medicinal plants after the markets in KwaZulu-Natal, and the ethnic diversity of the region's traders, healers, gatherers and consumers is influential in determining the traded floristic diversity and sources of the plants harvested for the multinational trade (Williams et al. 2000).

The trade is differentiated into two sectors, namely formal business and informal street markets (Williams et al. 1997). Traders, including traditional healers, selling traditional medicines from premises called '*muti*' shops, represent the formal sector. In 1994, there were estimated to be 244 shops in the region selling traditional medicines (Williams et al. 1997), the majority of which were owned by Black traditional healers (52%) and Indian merchants (25%).

Commercial gatherers and traders selling plants from the pavements and street markets, on the other hand, represent the informal sector. Located in Johannesburg, the 'Faraday Street' market is the Witwatersrand's only informal wholesale and retail street market for traditional medicine. Ninety-seven percent of the approximately 166 traders are migrants to the Witwatersrand, of whom 90% regard the province of KwaZulu-Natal as "home" (Williams 2003). Customers to the market are primarily traditional healers from Gauteng townships, owners of *muti* shops, and sometimes patients seeking treatment from the traders that are traditional healers. Most of the traders earn less than R100 a week (US\$1 $\approx$ R6.51, June 2004) (Williams 2003).

Methods

Market surveys

Between February 1994 and January 2001, two semi-quantitative surveys of plants traded for traditional medicine were conducted within the Witwatersrand. The first survey in 1994, based on a stratified random sample of 50 research participants from 50 *muti* shops, appraised the nature of the plant trade in the formal sector. The second survey in 2001, a stratified random survey of 100 street traders in the Faraday Street informal market, was conducted on behalf of the provincial Directorate for Nature Conservation for Gauteng. The surveys were based on questionnaires and structured interviews, and an inventory of all common names of plants sold by each trader was compiled. The characteristics of the trade and the species sold within the shops and at the market have already been quantitatively and qualitatively described (Williams et al. 2000, 2001; Williams 2003).

Synthesis of plant inventories

Identification of the species traded was mainly achieved by matching vernacular names to botanical names from previously published studies that, for the most part, are reliable because of the credible body of literature existing for ethnobotanical names in South Africa. In some cases, species were visually identifiable or were identified later from purchased specimens. Species identification through published records is problematic, and errors in identification are likely to have occurred. However, this was considered the most expedient mechanism for identifying the large numbers of inventoried species sold by each of the traders surveyed. In order to make a distinction between vernacular names synonymous with >1 species, analysis by '*ethnospecies*' instead of by botanical species was used. '*Ethnospecies*' (Hanazaki et al. 2000) takes into account the folk or common name given to one or several species quoted during interviews. The Zulu name '*iNgwavuma*', for example, is the ethnospecies name designating *Elaeodendron transvaalense* (Burtt Davy) R.H. Archer, whilst the ethnospecies '*iMphepho*' applies to at least six species of *Helichrysum*. *iMphepho* was cited 17 times during the Faraday market survey, however, only one of the six potential *Helichrysum* species would have been sold at each stall and the most prevalent species is not known. Therefore, wherever appropriate, the data were quantified based on the number and frequency of occurrence of ethnospecies to avoid repetitions and hence any bias/inaccuracies in reporting the results. In this paper, all citations of *species* are citations for plant *ethnospecies*.

Six data sets derived from the Witwatersrand medicinal plant trade were subject to analysis with the 22 diversity measures. This was to evaluate the effect of different sample sizes, trader ethnicities and formal/informal market sectors in the appraisal of patterns in the utilisation and trade of traditional plant medicines. The sample of 50 *muti* shops (abbreviated as *MS: All 1994* in the graphs) was subdivided into three smaller subsamples based on the ethnicity of the shop owner, namely Black ($n = 28$ shops; *MS: Black 1994*), Indian ($n = 20$ shops; *MS: Indian 1994*) and White ($n = 2$ shops; *MS: White 1994*). The Faraday market data (*ST: Faraday*) were not subdivided for the initial appraisal of the indices. An earlier survey conducted by the author in 1992 of seven *muti* shops (*MS: 1992*) was included to compare the effect of sample size (Williams 1992). Whenever appropriate, the performance of the indices was compared with a seventh data set – a sample of 17 informal vendors derived from an inventory compiled for medicinal plants traded on the western boundary of the Kruger National Park, South Africa (Botha 2001; Botha et al. 2001). The dataset is abbreviated as '*ST: WBKP*' in the graphs. Later, by way of an independent example comparing intra-sample diversity for selected indices, the Faraday data matrix was subdivided into 'healer' and 'non-healer' traders.

Calculating indices

The calculation of an index to evaluate ecological diversity relies on information regarding the *number* and *frequency of occurrence* of species in a sampled community. In order to calculate diversity indices for ethnobotanical purposes it is necessary to have data on the number of *individual informants* (e.g., resource users/traders) who *cited* the plant species (Begossi 1996). Since the inventory of plants sold by each trader in this study recorded the presence of an ethnospecies, each ethnospecies was recorded/cited once per trader and, therefore, incidence/occurrence equals abundance of the ethnospecies per trader sample/inventory.

References for the formulae and software used to calculate the indices are listed in Table 1. The statistical distributions used to fit species abundance observations may be used for fitting species occurrences (Hayek and Buzas 1997). N occurrences may be substituted for N individuals (Hayek and Buzas 1997). In the calculations, $n =$ number of samples (e.g., number of *muti* shops or street traders surveyed) and $N =$ number of citations or occurrences of ethnospecies.

The species accumulation curves and cumulative diversity curves were constructed from variables and diversity statistics computed by *EstimateS* (Colwell 2001). In cases where *EstimateS* did not directly compute the diversity measure (e.g., Margalef and Menhinick's indices, Hill's diversity numbers N1 and N2, evenness measure E1–E5), then the formulae cited in Magurran (1988) and the appropriate variables computed by *EstimateS* were used to calculate the indices and construct the graphs. It is important to note that *EstimateS* computes the

Table 1. References for methods used to compute the indices and measures applied to the data.

Index/Measure	Reference
Species richness indices	
# Species (S or N0)	Discerned from observation of the data set
Margalef (D_{Mg})	Ludwig and Reynolds (1989) – program 'SPDIVERS.BAS' or Magurran (1988) and Colwell (2001)*
Menhinick (D_{Mn})	Ludwig and Reynolds (1989) – program 'SPDIVERS.BAS' or Magurran (1988) and Colwell (2001)*
Diversity indices	
Shannon–Wiener (H')	Ludwig and Reynolds (1989) – program 'SPDIVERS.BAS' or Colwell (2001)*
Shannon–Wiener (H_{max})	Magurran (1988)
Comparing Shannon indices	Zar (1984); Magurran (1988); Murali et al. (1996) or use standard deviation for H' from Colwell (2001) *
Brillouin measure (HB)	Zar (1984); Magurran (1988); Krebs (1989)
Simpson (λ)	Ludwig and Reynolds (1989) – program 'SPDIVERS.BAS' or Colwell (2001)*
Simpson ($-\ln \lambda$)	Pielou (1975); Colwell (2001)*
Berger–Parker (d)	Magurran (1988)
McIntosh's dominance (D)	Magurran (1988)
Fisher's alpha (α)	Magurran (1988); Krebs (1989); Colwell (2001)*
Hill's Diversity Number N1	Ludwig and Reynolds (1989) – program 'SPDIVERS.BAS' or Magurran (1988) and Colwell (2001)*
Hill's Diversity Number N2	Ludwig and Reynolds (1989) – program 'SPDIVERS.BAS' or Magurran (1988) and Colwell (2001)*
Hill's Diversity Number $N\infty$	Magurran (1988)
Evenness measures	
Shannon (J′ or E1) (or, Pielou's J)	Ludwig and Reynolds (1989) – program 'SPDIVERS.BAS' or Magurran (1988) and Colwell (2001)*
E2 (or, Buzas′ & Gibson's E)	Ludwig and Reynolds (1989) – program 'SPDIVERS.BAS' or Magurran (1988) and Colwell (2001)*
E3, E4 and E5	Ludwig and Reynolds (1989) – program 'SPDIVERS.BAS' or Magurran (1988) and Colwell (2001)*
Brillouin (J)	Krebs (1989)
McIntosh's (E)	Magurran (1988)

*Indicates that *EstimateS* (Colwell 2001) either computes the index and/or the parameters that can be inserted into the equations obtained from the other references listed.

reciprocal of Simpson's λ. In the formulae for diversity indices, any logarithmic base may be used (Zar 1984). As a way of standardising the results, the natural log (ln) was used throughout.

Results and discussion

Species richness

Numerical species richness (S), or the *number of species* in a sample of a specified size, is an instantly comprehensible expression of species diversity

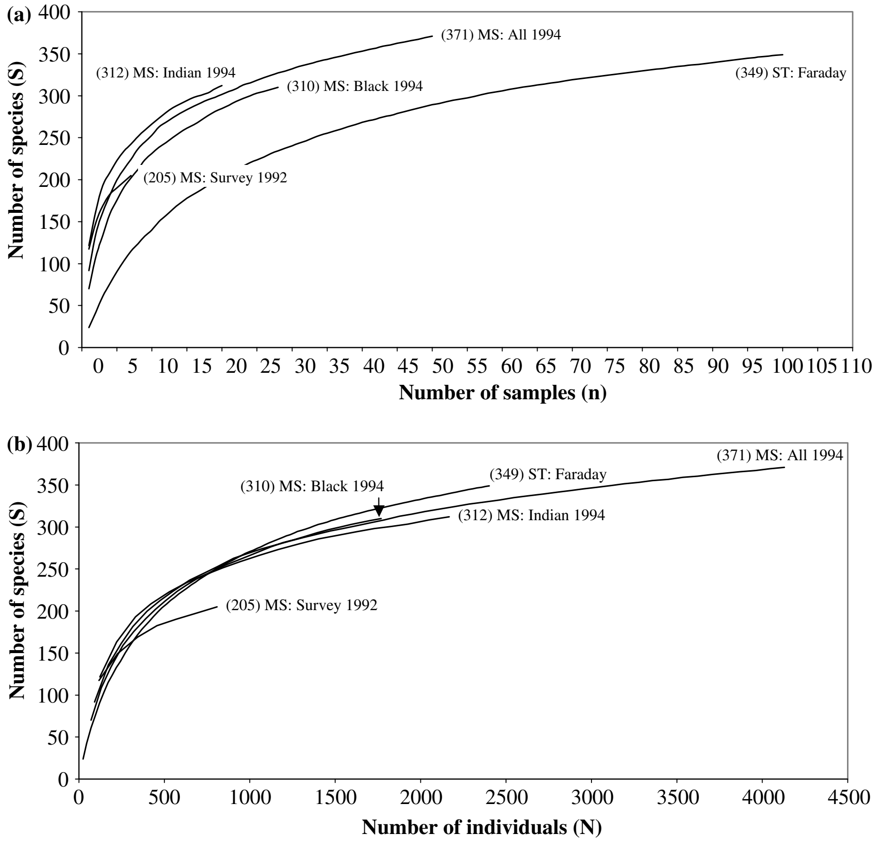


Figure 1. Species accumulation curves (or collector's curves) for plant ethnospecies traded as traditional medicine in muti shops and in a street market in the Witwatersrand. The curves represent successively pooled and randomly ordered samples (a) and individuals (b). The curves were computed using *EstimateS* (Colwell 2001). The total number of ethnospecies per sample (*S*) is labelled in brackets at the end of each curve. (MS: *Muti shops*; ST: Street traders).

(Magurran 1988). *S* is related to the total number of individuals (*N*) summed over all *S* species recorded. As sampling effort increases, more individuals are encountered and more species are likely to be recorded (Hayek and Buzas 1997). The relative abundance of species, however, is important and a number of simple indices have been derived using some combination of *S* and *N*. These indices include Margalef's and Menhinick's index of species richness. While these indices are easy to calculate, they are (like *S*) sensitive to sample size.

The relationship between *S* and *N* may be viewed by plotting a *species accumulation curve* (Figure 1a, b), also termed a *species effort curve* or *collector's curve* – so called because the cumulative number of species is plotted against some measure of the effort it took to obtain that sample of species (Hayek and Buzas 1997). Compared to interpreting the single numerical value of species richness for the randomly pooled samples (Table 2), plotting the curves facilitates improved interpretation of species richness results for different samples of varying size. Comparing raw taxon counts (and index values) for two or more assemblages/samples will quite generally produce misleading results (Gotelli and Colwell 2001). Differences in measured species richness between communities may be the result of differences in underlying species richness, differences in the shape of the relative abundance distribution, or because of differences in the number of individuals counted (Gotelli and Colwell 2001).

Whereas fewer *muti* shop traders were sampled compared to street traders (n) (Figure 1a), the number of individual plants (N) recorded in the shops was greater (Figure 1b), and hence the numerical richness per trader is greater for the *muti* shops. There is also a similarity in shape and clustering of curves for the shop data ('*MS*'), even the 1992 survey of seven shops, making them distinct from the curve of the street trader ('*ST*') sample (Figure 1a). The initial steep gradients of the curves for the shop data show that more ethnospecies per shop are sold (mean = 83) and consequently the accumulation of ethnospecies is more rapid, even for smaller sample sizes. Street traders, by contrast, sell fewer ethnospecies per trader (mean = 24) and consequently the accumulation of ethnospecies is slower. When samples are highly variable in terms of plant diversity amongst traders, then more samples are needed to fully represent the trade in medicinal plants, while if the samples show little variation then fewer traders need to be sampled.

The Margalef and Menhinick indices have been cited as being inadequate by several authors (e.g., Brower and Zar 1977; Magurran 1988; Hayek and Buzas 1997) because the indices lack the ability to differentiate the species richness of samples having similar S and N. Looking at Table 2, the performance of these indices as a single numerical value for pooled samples is difficult to adequately judge. However, plotting the performance of an index as samples are successively pooled and individuals are accumulated is a useful procedure for aiding the interpretation of plant availability within the different trader groups (Figures 2 and 3).

The graphs of Margalef's index show how species richness increases until eventually the curve levels off with increasing sample size and the number of individuals inventoried (Figure 2a, b). The point at which the curve flattens indicates a minimum viable sample size on which a diversity or richness index should be based (Magurran 1988). The curve of S (Figure 1) can also be constructed for this purpose i.e., to determine the minimum requisite sampling effort. The sample of species sold by White shop traders ($n = 2$) is numerically inadequate and only two points of richness could be plotted on the graphs. However, the sampling strategy for the 50 traders selected for the shop survey (including the 2 White traders) was stratified random, and therefore trader ethnicities were proportionately representative within the sample to minimize participation biases (Williams et al. 1997). Sampling more White traders would only have been necessary if it had been an important aim to compare the shops of different trader classes, but it would have biased the overall results of the

Table 2. Species richness indices calculated for six data sets sampled from Witwatersrand traders of traditional medicine. n = the number of samples (traders and/or shops surveyed); N = number of individual ethnospecies observed; S = number of ethnospecies counted.

Index/measure	1992 *Muti* Shop Survey (n = 7) N = 809	1994 *Muti* Shop Survey (N = 50)				2001 Street trader survey (n = 100) N = 2402
		'White-owned' (n = 2) N = 193	'Indian-owned' (n = 20) N = 2168	'Black-owned' (n = 28) N = 1769	Total ('All') shops (n = 50) N = 4129	
# Species (S or N0 or $e^{H_{max}}$)	205	144	313	310	371	349
Margalef (D_{Mg})	30.5	27.2	40.5	41.3	44.4	44.7
Menhinick (D_{Mn})	7.2	10.4	6.7	7.4	5.8	7.1

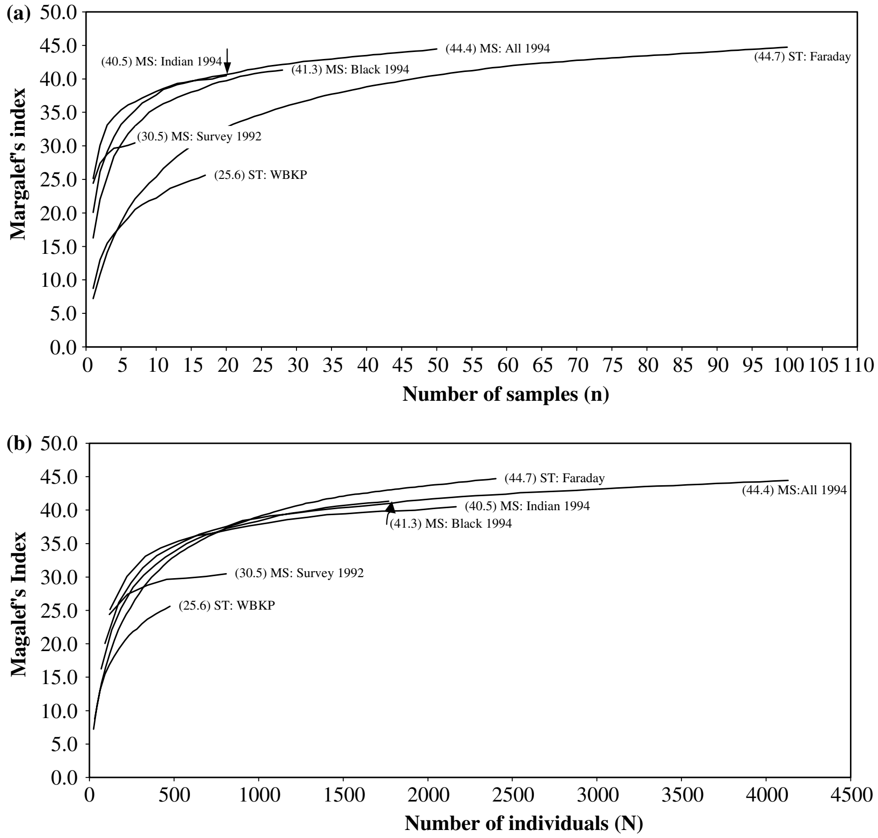


Figure 2. The cumulative species richness curves of Margalef's index for ethnospecies traded on the Witwatersrand for both samples (a) and individuals (b). 'WBKP' is a sample of 17 informal vendors trading medicinal plants on the western boundary of the Kruger National Park, South Africa (Botha 2001). The overall value of the index for the randomly pooled samples is labelled in brackets at the end of each curve. The formula for the index is $D_{\mathrm{Mg}} = (S{-}1)/\ln N$. (MS: *Muti* shops; ST: Street traders).

study. The minimum viable sample size (i.e., the number of research participants) necessary for assessing species richness is larger for informal street traders than for shop traders. The evidence for this is reflected in the species accumulation and diversity graphs for street traders, which show that a larger sampling effort is necessary before the curves begin to reach an asymptote (Figures 1 and 2).

Evidence for the distinctive trading patterns in species richness in formal and informal trading sectors are supported by the sample of 17 informal vendors from the western boundary of the Kruger National Park (*ST: WBKP*, Figure 2a). Despite the small sample size (which was conducted with 73% of the vendors at the site), the curve adopts a similar shape to the Faraday Street trader sample (*ST: Faraday*) and exhibits low ethnospecies richness per trader.

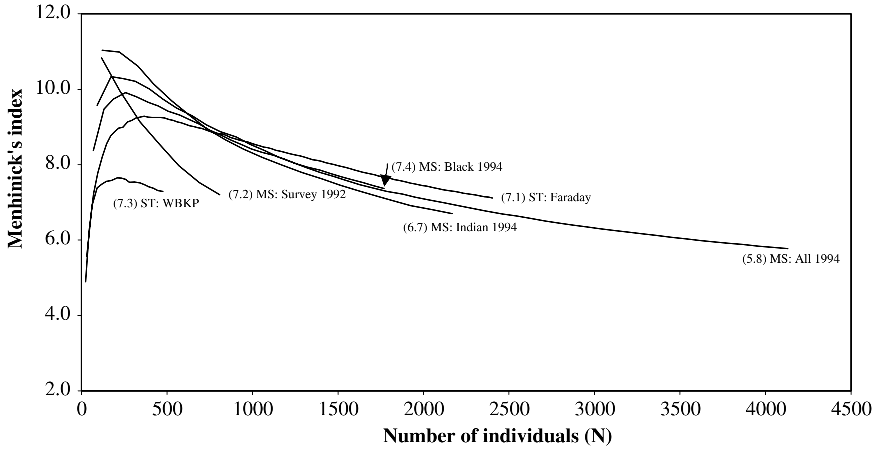


Figure 3. The cumulative species richness curves of Menhinick's index for ethnospecies traded on the Witwatersrand. The overall value of the index for the randomly pooled samples is labelled at the end of each curve. The formula for the index is $D_{Mn} = S/\sqrt{N}$. (MS: *Muti* shops; ST: Street traders; WBKP: Western Boundary Kruger National Park).

Gotelli and Colwell (2001) recommend that when *comparing species richness between samples*, the number of taxa should be plotted as a function of the accumulated number of *individuals* (not *samples*) because datasets may differ systematically in the mean number of individuals per sample. Figure 2b, therefore, shows that the relative abundance and richness of ethnospecies sold at the Faraday street market is higher than that of the shops. Margalef's index for Faraday is 44.7 compared to 44.4 for 'All shops' (*MS: All*). The relative abundance and richness of ethnospecies sold by Black traders is slightly higher than that of Indian traders.

The graph of Menhinick's index (Figure 3) corroborates earlier evidence derived from Margalef's index that street traders keep a lower species richness per trader, and therefore increased sampling effort is required for the curve to reach a peak before declining as n and N increase. The *ST: WBKP* sample shows a similar pattern to the Faraday Street trader data, which is different to the pattern shown by the *muti* shops. The numerical richness values for the index similarly indicate the Faraday sample to be relatively more species rich than the 'All shops', and the Black trader sample to be more species rich than the Indian trader sample.

Species diversity or heterogeneity

Indices of diversity or heterogeneity incorporate both richness and evenness into a single value and are based on *the proportional abundance* of species in a sample (Ludwig and Reynolds 1988; Magurran 1988). These measures are attractive to researchers because, unlike the species abundance models, no assumptions are made about the underlying distributions of the data (Hayek and Buzas 1997). There are four categories of indices. First are measures derived from *information theory* (e.g., Shannon–Wiener and Brillouin), based on the rationale that diversity or information in a natural system may be measured in a similar way to information contained in a code or message (Magurran 1988). The second category of indices are the *dominance* indices (e.g., Simpson, McIntosh and Berger–Parker), so called because they are weighted towards the abundances of the commonest species (Magurran 1988). A third category of diversity index is '*Hill's diversity numbers*'. The numbers, developed by Hill (1973), show how diversity indices are mathematically related and may be arranged in a sequence depending on whether they measure species richness (weighted towards uncommon species) or dominance (weighted towards abundant species) (Magurran 1988). Interpreting the single statistic for each index can be problematic (Table 3). A fourth category of diversity index is derived from the logarithmic series abundance model, namely Fisher's alpha.

Information theory indices

The Shannon index (H') measures the average degree of ''uncertainty'' in predicting to what species individuals chosen at random will belong (Ludwig and Reynolds 1988). Uncertainty may be visualised as being synonymous with diversity (Krebs 1989), therefore, the higher the degree of uncertainty, the higher the diversity and the greater the degree of uncertainty in correctly predicting the identity of the next species chosen at random. The average uncertainty (H') increases as S increases, as seen in Figure 4 when compared with Figure 1. Figure 4a shows that there is a distinction in the curves between formal and informal traders (Figure 4a), with the overall degree of uncertainty and diversity being higher in the formal sector (Figure 4b). The higher predictability of a species' identity in the street markets is because of the lower mean species richness per trader, as discussed in the previous section. Informal traders sell fewer species and one may more comfortably predict what popular species most traders are likely to sell. In terms of determining the minimum sample size necessary to assess the Shannon index (as indicated by the point at which the curve levels off), the curve for shop traders begins to reach an asymptote at around 15–25 samples, compared to 25–30 for the street traders.

Maximum uncertainty (H_{max}) will occur when each species in a sample is equally represented (Table 3) (Hayek and Buzas 1997). The more species there are, the more evenly the individuals are spread across the species, the higher will be the value of H' because there will be greater uncertainty as to which species will most likely be observed next time they are chosen at random. It appears that a characteristic of ethnobotanical samples (especially those of large sample sizes) is that H' is high. In examples described in Magurran (1988) for ''natural'' communities (e.g., diversity of birds in woodlands; species diversity in plantations etc.,) H' ranged between 1.38 and 3.54. By contrast, Begossi (1996) estimated H' to be between 2.99 and 5.95 (average = 4.6) for

Table 3. Species diversity indices calculated for six data sets sampled from Witwatersrand traders of traditional plant medicine.

Index/Measure	1992 *Muti* Shop Survey ($n = 7$) $N = 809$	1994 *Muti* Shop Survey ($N = 50$)				2001 Street trader survey ($n = 100$) $N = 2402$
		'White-owned' ($n = 2$) $N = 193$	'Indian-owned' ($n = 20$) $N = 2168$	'Black-owned' ($n = 28$) $N = 1769$	Total ('All') shops ($n = 50$) $N = 4129$	
Shannon-Wiener (H')	5.16	4.91	5.43	5.38	5.46	5.33
Shannon-Wiener (H_{max})	5.32	4.97	5.75	5.74	5.92	5.86
Brillouin measure (HB)	2.08	1.78	2.25	2.22	2.30	2.21
Simpson (λ)	0.0050	0.0026	0.0047	0.0051	0.0049	0.0066
Simpson ($-\ln \lambda$)	5.29	5.95	5.36	5.28	5.32	5.02
Berger-Parker (d)	0.0087	0.0104	0.0088	0.0113	0.0094	0.0249
McIntosh's dominance (D)	0.954	0.982	0.949	0.947	0.942	0.935
Fisher's alpha (α)	88.49	257.28	99.95	108.85	98.75	112.26
Hill's Diversity Number N1($e^{H'}$)*	174.9	135.7	227.5	216.8	233.9	206.5
Hill's Diversity Number N2($1/\lambda$)*	198.9	378.1	213.2	195.8	200.1	150.9
Hill's Diversity Number N∞*	115.6	96.5	114.1	88.4	105.9	40.0

*The result is a count of the number of ethnospecies.

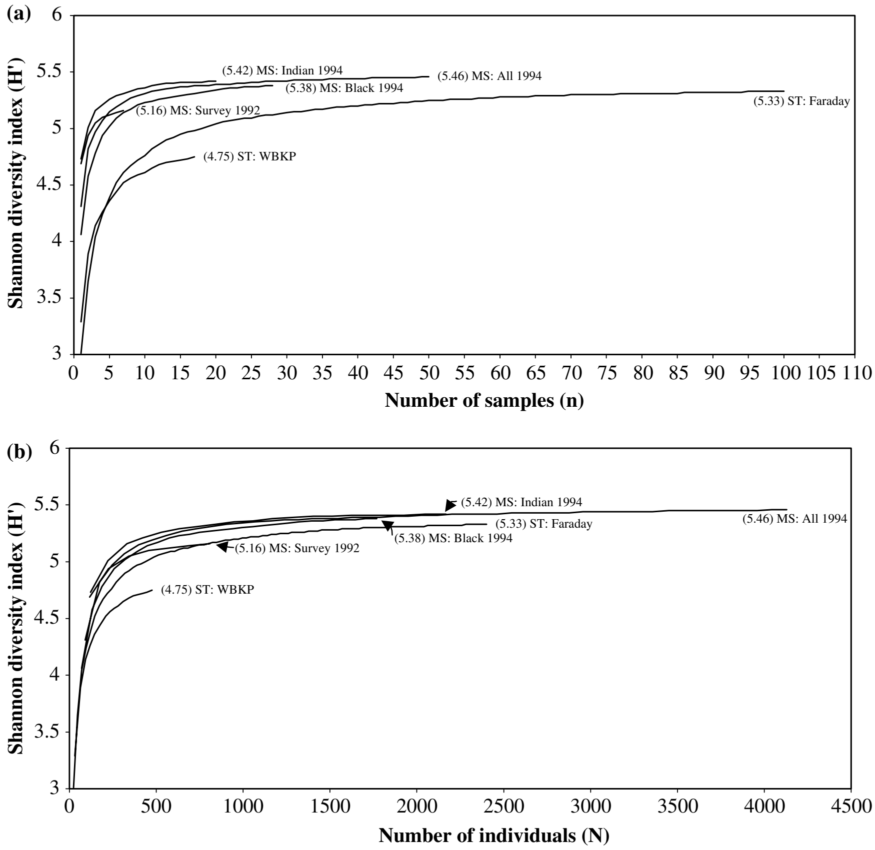


Figure 4. The cumulative diversity curves of Shannon's diversity index (H') for ethnospecies traded on the Witwatersrand for samples (a) and individuals (b). The overall value of the index for the randomly pooled samples is labelled in brackets at the end of each curve. (MS: *Muti* shops; ST: Street traders; WBKP: Western Boundary Kruger National Park).

eight ethnobotanical samples from mainly South America. In the South African study, H' ranged between 4.91 and 5.46 (Table 3).

When the Shannon index is obtained for two or more samples it is possible to test the null hypothesis that the diversities of the samples are equal (Zar 1984). The differences in the index between the samples is mostly significant at $p < 0.000001$, except for three comparisons which are approaching an equal diversity. The diversity of plants sold by Black traders when compared with both the diversities of the Faraday traders and Indian-owned shops is $p = 0.0021$. The least significantly different samples are those of Indian traders and All shops, where $p = 0.0017$ – this would suggest that the characteristics and diversity of All Shops ($n = 50$) is largely due to the influence of the sample of Indian traders ($n = 20$) within it.

The Brillouin index (HB) is similar to Shannon, and the use of this index instead of Shannon is recommended when randomness of a sample cannot be guaranteed (Magurran 1988). The values for HB are lower than H' (Table 3) and show similar numerical rankings for species diversity between samples. The obstacle to using this index is the calculation of very large factorials; additionally, it was not possible to derive the accumulated diversity curve. While many authors recommend the use of HB over H' (e.g., Magurran 1988), the simplicity of calculating Shannon is to its advantage and has led to its widespread acceptability as an index – however, Shannon is sensitive to sample size, thus indices such as Simpson's and Fisher's are sometimes preferred by researchers. Magurran (1988) says that "ideally (Fisher's) alpha should replace the Shannon index as the preferred measure".

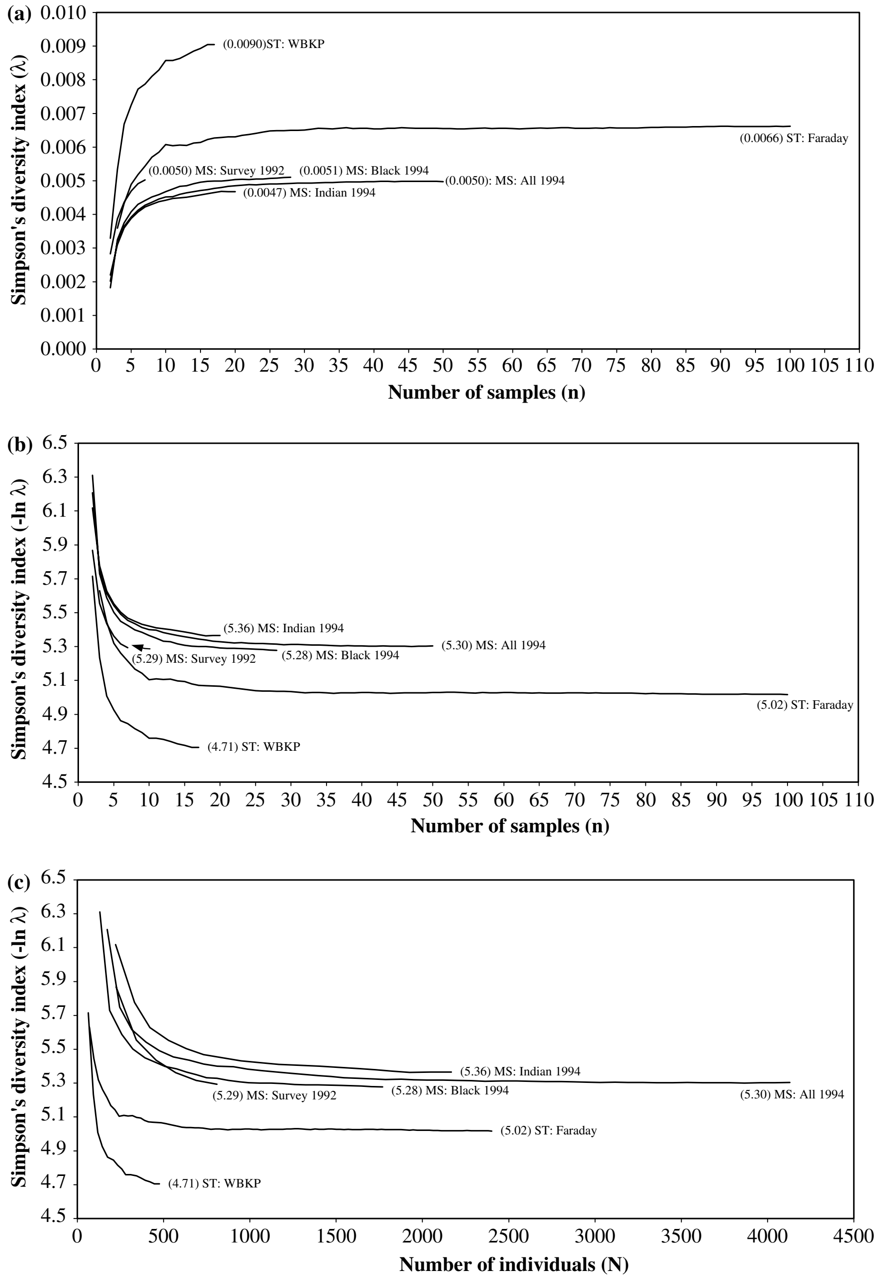
Dominance indices

Simpson's index (λ) proposes that diversity is inversely related to the probability that two individuals picked at random from a sample belong to the same species. Simply stated, if the probability (λ) is high (λ approaches 1) that both individuals belong to the *same* species, then the diversity of the sample is low and *vice versa* (Pielou 1975; Ludwig and Reynolds 1988; Krebs 1989) (e.g., Figure 5a). However, because λ decreases as diversity increases, Simpson's index is usually expressed as $1 - \lambda$ (the probability that two individuals chosen at random are *different species*) or $1/\lambda$ (also known as Hill's number N2, which is a measure of the number of very abundant species in a sample). A rarely cited function of λ is $-\ln \lambda$ (Pielou 1975) (Figure 5b, c), and is preferred by the authors as a diversity indicator. The function does not represent a probability, but a single diversity statistic that increases as diversity increases.

Indian shop traders, followed by the total sample of shops (*MS: All)* and Black traders, have the largest diversity of plants for sale, especially compared to the street traders (Figure 5c). As with the other indices discussed so far, there is a distinction between the diversity of plants for sale by the *muti* shops and street traders. There is, therefore, a higher probability that two plants selected at random from different street trader stalls belong to the same species than for two plants selected from different *muti* shops (Figure 5a). As a result, there is a greater dominance of certain species sold by the street traders.

When $\lambda = 1$, most individuals from a sample are concentrated in a single species, and therefore dominance of species within the sample is high. Values for λ for the ethnobotanical samples investigated in this paper are relatively low ($\lambda < 0.008$), and therefore the overall dominance of species is relatively low and diversity is relatively high. By comparison, Hanazaki et al. (2000) calculated λ to be between 0.015 and 0.06 for plants used by native inhabitants from the Atlantic Forest coast in south-eastern Brazil. The overall diversity of plants used by the community investigated in Brazil is, therefore, lower and dominance of plants is higher when compared with the South African study.

The minimum sample size necessary for evaluating Simpson's index is between 15–20 for street traders and 20+ for *muti* shops (Figure 5b). The graph



shows that the sample '*MS: Survey 1992*' is too small to assess the index, and therefore comparisons of diversity with the other samples. Additionally, curves for the street traders show that despite the smaller number of individuals sold

◀

Figure 5. The cumulative diversity curves of Simpson's diversity index (λ) for ethnospecies traded on the Witwatersrand for samples (a, b) and individuals (c). (a) plots the standard form of the index, namely λ, the probability that two individuals will belong to the same species. b and c plot $-\ln \lambda$, a rarely cited form of the index recommended by Pielou (1975) that expresses λ as a diversity statistic that increases as diversity increases. The overall value of the index for the randomly pooled samples is labelled in brackets at the end of each curve. (MS: *Muti* shops; ST: Street traders; WBKP: Western Boundary Kruger National Park).

per street trader, that dominance of individuals related to their relative abundance is established sooner (indicated by the point at which the curve levels off). This confirms the Shannon index results, i.e., that the probability of encountering the same species is higher between street traders. Additionally, there are fewer abundant species sold by the street traders compared to the *muti* shops (as demonstrated by the curve for N2 [or $1/\lambda$] in Figure 6).

The Berger–Parker index is a dominance measure that expresses the proportional abundance of the most abundant species ($d = N_{max}/N$) (Table 3). The index is independent of S, but is subject to bias caused by fluctuations in the abundance of the commonest species (Magurran 1988). Like Simpson's index (λ), diversity increases and dominance decreases as d decreases. The results are concordant with Shannon and Simpson's index, except for the '*MS: Survey 1992*' sample which shows a higher diversity than for the Indian traders. However, it was established in Figure 5a that the sample size was too small in this sample to assess species dominance and hence the result is disregarded in

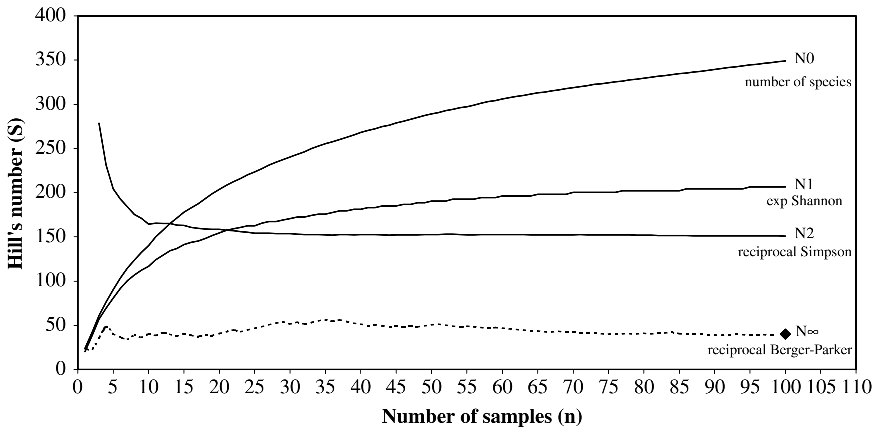


Figure 6. The cumulative diversity curves for Hill's numbers (N0, N1, N2 and $N\infty$) for ethnospecies traded in the Faraday Street market. The curves for N0, N1, and N2 were based on successively pooled and randomly ordered samples, whilst $N\infty$ has been estimated from non-randomised values directly from the dataset because the requisite parameters for a randomised and cumulative curve were not available on the software. The final value for $N\infty$ is shown as ◇ because it does not depend on the order of the randomised data matrix. The unit on the *y*-axis is number of plant ethnospecies.

this instance. It was not possible to plot the curves for the performance of this index as *EstimateS* does not calculate the requisite parameters.

McIntosh's index (D) is a measure of diversity independent of N. As dominance increases (related to the increase in abundance of species in the sample), the value of D appears to decrease. However, the discriminant ability of this index in samples of different sizes is poor and, coupled with the inability to graph the index, the performance and usefulness of this index was difficult to evaluate.

Hill's diversity numbers

Hill's numbers are the easiest diversity statistic to interpret. The diversity numbers are in units of *number of species* and measure what Hill calls the *effective number of species* present in a sample (Hill 1973; Ludwig and Reynolds 1988). The numbers are mathematically related to the Shannon, Simpson and Berger–Parker indices (Table 3). As the number of species (N0 or S) increases, less weight is placed on 'rare' species, and lower values are obtained for N1, N2 and N∞ since they measure the number of *abundant*, *very abundant* and *most abundant* species in a sample respectively (Table 3). 'Rare' species in an ethnobotanical context are those species with low incidences/abundances in the sample (i.e., low values of n).

Hill's numbers may be plotted, for example the Faraday Street trader sample in Figure 6. Like the other indices previously described, the point at which the curve levels off gives an indication of the minimum viable sample size needed to assess the index. The minimum sample size required for the street traders is around 30 (Figure 6), compared to $\pm$20 for shop traders (not shown graphically). The distinctive high values for the first part of the N2 curve (derived from Simpson's index) is because N2 is weighted in favour of the commonest species. Because incidence equals abundance in these samples, adding new or 'rare' species to the sample (as n or N increases), decreases the relative abundance of the commonest species until the value stabilises when fewer new species are added. N1, on the other hand, shows a steady increase until the curve reaches an asymptote – this is because the function is derived from the Shannon index, and therefore weighted in favour of species richness. As the number of new species increases, the value of the curve increases until it eventually levels off when very few new species are added to the sample as n increases.

It was not possible to plot N∞ for the randomised and successively pooled samples because the requisite parameters were not available on the *EstimateS* software. Except for the final value of N∞ (shown as ◆ in Figure 6 for the total pooled sample), the projection of the curve has been estimated from non-randomised values from the data matrix for Faraday. The curves for the *muti* shop samples are similar in shape to the street trader data, but higher in value. The curves are not plotted because Table 3 adequately expresses the results.

Values for N1, N2 and N∞ are higher for the *muti* shops than for the Faraday Street market (Table 3), indicating that if species abundance equates with plant popularity, then there is greater dominance of a few popular (abundant) species in the street market. In other words, the street traders sell a smaller number of ethnospecies that have very high occurrences within the market. *Muti* shops, on the other hand, sell a larger range of species with equally high abundances.

This pattern is related to factors already discussed in this paper, i.e., that street traders have neither the space nor the financial capacity to sell large numbers of species. They, therefore, sell a smaller range of ethnospecies that have assured commercial value and are likely to have a higher restocking potential. Additionally, if one trader does not have a certain plant that the customer is looking for, then there are at least 160+ other traders in the market that might sell the plant. *Muti* shop traders, on the other hand, have larger trading spaces and financial freedom and can afford to stock a large range of species in their shop – i.e., they are 'one-stop-shops'. The number of species represented by N2 and $N\infty$ are indicators of the number of ethnospecies within the markets that are candidates for more immediate conservation action, assuming that high incidence is an indicator of potential risk.

Fisher's alpha

Generated from a species abundance model, Fisher's alpha (α) is a constant used to fit the logarithmic series distribution model once the parameter x has been solved for iteratively. The index is also known as the log *series alpha* (Magurran 1988), and has been adjudicated as a good, if not one of the best, measures of species diversity by several authors (e.g., Magurran 1988; Krebs 1989; Hayek and Buzas 1997) even when the underlying species abundances do not follow a log series distribution.

Alpha is low when the number of species is low (Table 3, Figure 7), and therefore the smaller samples with fewer ethnospecies have smaller values of α (e.g., WBKP and MS: Survey 1992). The index is less affected by the abundance of the rarest or commonest species than either H' or λ, respectively (Magurran 1988), and depends more on the number of species of intermediate abundance. According to Hayek and Buzas (1997), Fisher's α is a number close to the number of species we expect to be represented by one individual – this would account for the high values of α in the initial part of the curves in Figure 7. Because the incidence of species sold at trader's shops/stalls equals abundance, the samples all initially have very high numbers of ethnospecies represented by one individual due to the nature of the sampling methods.

Hayek and Buzas (1997) recommend that α is used as a diversity measure when the parameter x of the log series model is $1 \geq x \geq 0.61$ because when $x \leq 0.61$ then $\alpha > S$ and the statistic becomes unacceptable and biologically meaningless (where $x = N/[N+\alpha]$). Greater confidence in the true value of α occurs when x is close to 1 and N is large. The point on the curves at which $x \geq 0.61$ is marked with *, and the final value of x is also shown (Figure 7). Eventhough α is a constant, it would appear from the data that α increases with N and S. Hayek and Buzas (1997) explain this apparent paradox in the

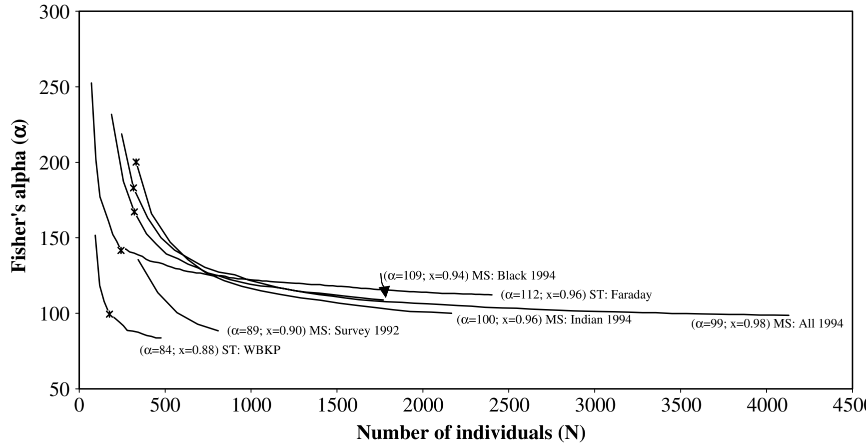


Figure 7. The cumulative diversity curves of Fisher's a for ethnospecies traded in the Witwatersrand. The overall value of α and x (a parameter required to fit the log series model) for the randomly pooled samples is labelled in brackets at the end of each curve. The point marked $*$ on the curves is where $x \leq 0.61$. (MS: *Muti* shops; ST: Street traders; WBKP: Western Boundary Kruger National Park).

following way: if the data fits the log series model *exactly*, then α is a constant independent of N. However, data rarely fits the models exactly in ''Mother nature'', and therefore an increase in α is observed as S and N accumulate. The results in Figure 7 show the Faraday Street traders to have greater species diversity than the other samples.

Evenness

Measures of evenness (or equitability) attempt to quantify the unequal representation of species against a hypothetical sample in which all species are equally abundant (Krebs 1989), i.e., the ratio of observed diversity to maximum possible diversity. Hence, evenness may be referred to as *relative diversity* or *homogeneity* (Brower and Zar 1977; Zar 1984). A low evenness means a high dominance in the use (or presence) of a few species (Begossi 1996). When all species are equally abundant, an evenness index would be at a maximum (of 1.0) and decrease towards zero as the relative abundances of the species diverge away from evenness (Ludwig and Reynolds 1988). Different measures of evenness have been proposed, most of which are expressions of Hill's numbers (Table 4).

All the indices gave different values but consistent rankings for the samples and subsamples (Table 4). Which index should, therefore, be chosen as a representative measure of how evenly distributed are the species sold by the Witwatersrand traders? Magurran (1988) recommends the use of the Brillouin evenness index, but the computation of very large factorials made it impossible to calculate. E1, also called the Shannon J′ or Pielou's J, is probably the most common evenness index in use but is strongly affected by species richness, and the addition of rare species (or singletons) can greatly change the value of E1 (Ludwig and Reynolds 1988). Hayek and Buzas (1997) recommend the use of E1 and E2 (also known as the Buzas and Gibson *E*). Ludwig and Reynolds (1988) further describe E3–E5, but consider E1–E3 to be of limited value because they are highly sensitive to the number of species in the sample. A general problem with all measures of evenness, however, is that they assume that the total number of species that could possibly be sampled is known (Krebs 1989). Since observed species numbers must always be less than true species richness, the evenness ratios are always *overestimated*, with the possible exception of E4 and E5.

E4 and E5 remain relatively constant with sampling variations and hence tend to be independent of sample size (Ludwig and Reynolds 1988). This is because E4 and E5 are computed as ratios where *S* is in both the numerator and the denominator, thus effectively cancelling the impact of the number of species in the sample (Ludwig and Reynolds 1988). However, E4 and E5 are not totally unaffected by the large number of singletons found in small samples, including the samples collected in the initial stages of research at a site before an adequate sample size is accumulated. Figure 8a shows how E5 > 1.0 until about sample 17 in the *muti* shops and about sample 22 in the street traders. An explanation for this feature of the index is as follows: initially N2 > N1 (Figure 6) because 'rare' species and singletons are to begin with very abundant in the ethnobotanical samples, thereafter declining in numbers as samples accumulate and the more dominant species become evident in the sample. This feature of the index is useful for determining the minimum viable sample size required for assessing evenness. E5 for samples *MS: White* and *MS: Survey 1992* is never less than 1.0 (Table 4), and therefore their evenness cannot be compared with the other samples. The results demonstrate that evenness is higher in the sample of Indian shop traders, followed by Black, All and Faraday Street, and therefore there is greater dominance in the sale of few species within the street market (Table 4). This result is consistent with the observations described earlier, namely that a high sample diversity means that it is more difficult to correctly predict a species chosen at random from a sample, and therefore the dominance of species is lower and evenness higher.

The values of E2 and E3, as well as E4 and E5, are similar, and therefore either may be used. However, Ludwig and Reynolds (1988) recommend the use of E5 as a measure of evenness because it is the least ambiguous. The authors also suggest calculating E1 because it is more widely used as a comparative index (e.g., Begossi 1996). The performance of E1 as an index is shown in Figure 8b for the sample of Faraday Street traders and shops (*MS: All*). The sensitivity of E1 to the addition of 'rare' species (singletons) is evident in the first part of the curve. When the Witwatersrand results are compared with six ethnobotanical samples from South America (Begossi 1996), it is evident that

Table 4. Species evenness indices calculated for six data sets sampled from Witwatersrand traders of traditional medicine.

Index/Measure	1992 *Muti* Shop Survey (n = 7) N = 809	1994 *Muti* Shop Survey (N = 50)				2001 Street Trader Survey (n = 100) N = 2402
		'White-owned' (n = 2) N = 193	'Indian-owned' (n = 20) N = 2168	'Black-owned' (n = 28) N = 1769	Total ('All') shops (n = 50) N = 4129	
E1 (Shannon J′*or Pielou's* J) (H'/H_{max})	0.970	0.988	0.945	0.938	0.922	0.910
E2 (N1/N0, or Buzas & Gibson E)	0.853	0.943	0.727	0.699	0.631	0.592
E3 (N1−1/N0−1)	0.853	0.942	0.726	0.698	0.629	0.590
E4 (N2/N1)	1.137	2.786	0.937	0.903	0.855	0.731
E5 (N2−1 / N1−1)	1.138	2.799	0.937	0.903	0.855	0.730
McIntosh's (E)	0.990	0.994	0.984	0.980	0.979	0.968
Brillouin (J)	Can't compute	Can't compute	Can't compute	Can't compute	Can't compute	Can't compute

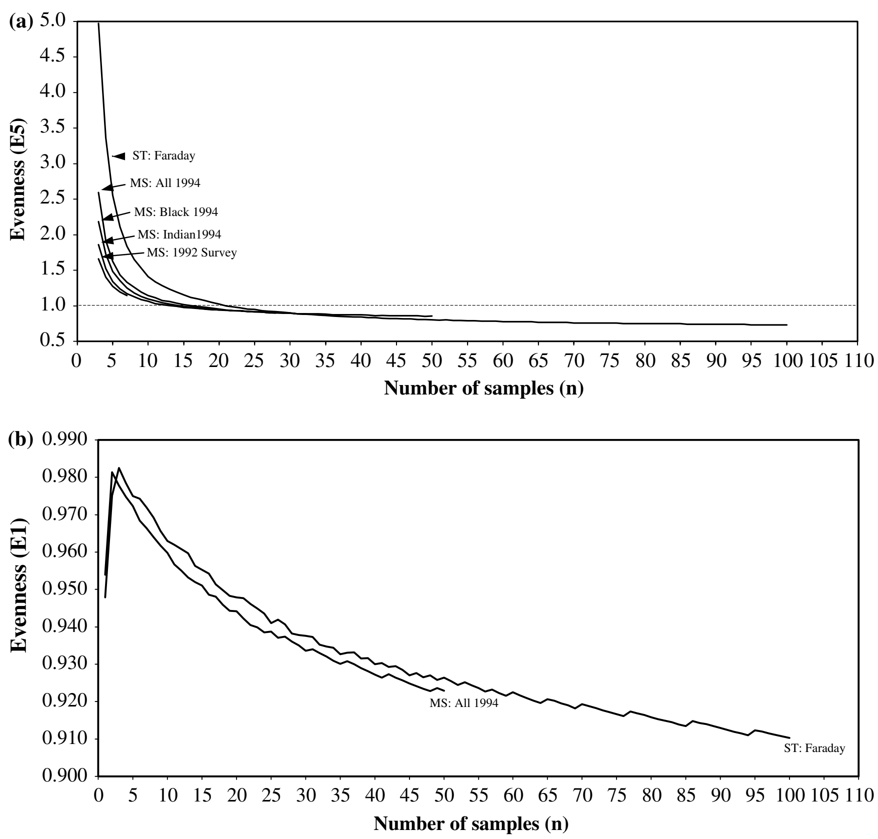


Figure 8. The cumulative evenness curves for (a) E5 and (b) E1 for ethnospecies traded in the Witwatersrand. (MS: *Muti* shops; ST: Street traders).

evenness is high overall for ethnobotanical samples (E1 > 0.90 on average), with very little overall dominance of certain species for use/sale. Values for E1 in the South American study range between 0.78 and 0.97 (average 0.91) compared to an average of 0.92 for the Witwatersrand traders.

Assessment of rare, intermediate and common ethnospecies

Indicator species are a useful adjunct to investigations of diversity (Magurran 1988). In ecology, they can provide an additional clue to how community structure is changing (Magurran 1988). In ethnobotany, indicator species are usually those in high demand by resource users and are at risk of over-exploitation and population decline. A question that frequently arises is: how does one objectively select criteria and categories for delimiting high risk species from those that are lower risk? Cunningham (2001) describes some of

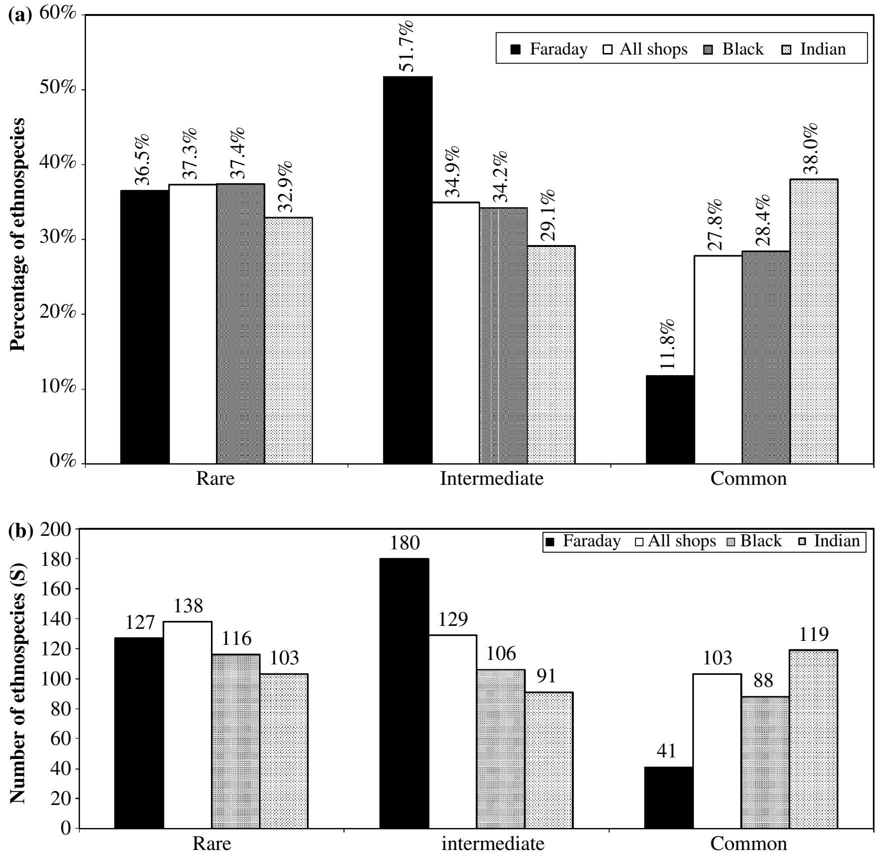


Figure 9. The percentage (a) and number (b) of species of rare, intermediate and common abundances within the Faraday street trader sample and the muti shop survey of 1994. The categories are derived from Hill's numbers N0, N1 and N∞.

the categories used for choosing priority species for monitoring as "filters" which help to sift out species that are likely to be more vulnerable to over-harvesting. While complex and comprehensive models and methods exist for 'filtering' species (the authors are currently addressing this question in more detail in forthcoming papers), a simple (albeit crude) method obtained from Hill's numbers can be used as a first step in the process of prioritising species for monitoring.

Hill's numbers N1, N2 and N∞ measure the number of *abundant, very abundant* and *most abundant* species in a sample, respectively. In addition, N0 equals S – the total number of species. The number of species of *rare, intermediate* and *common* abundances within a sample were defined in the following way: common $\approx$ N∞; intermediate $\approx$ N1–N∞; and rare $\approx$ N0–N1. 'Rare' species, as previously mentioned, are those species with low incidences/abun-

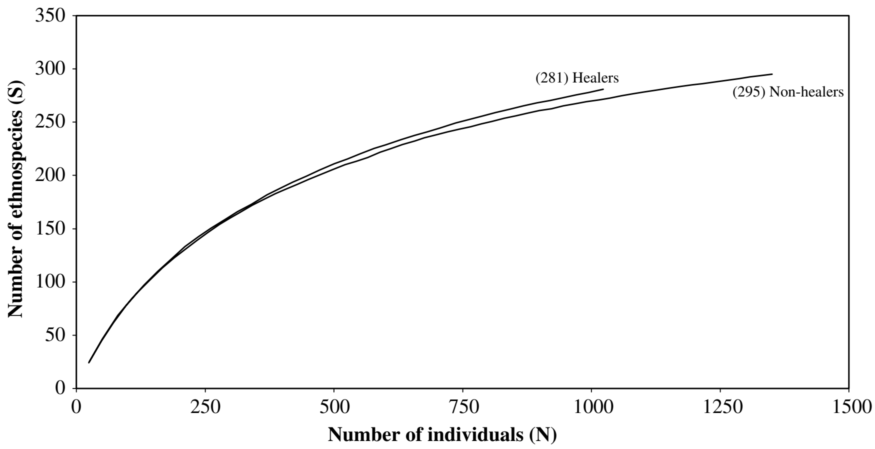


Figure 10. Species accumulation curves for ethnospecies sold by healer and non-healer traders in the Faraday street market.

dances in the sample (i.e., low values of n) and are not necessarily endangered. Rare species are, therefore, the remaining species in a sample that are not abundant (N1). Because N1 and N2 represent species present in abundance in a sample (but not of 'most' abundance like $N\infty$), they were combined to produce the category 'intermediate'.

Figure 9 shows the number and percentage of ethnospecies categorised as rare, intermediate and common in abundance from the Faraday street market survey and the *muti* shop survey of 1994. The purpose of this paper, is not to discuss what those species actually are, but to derive a method to more easily delimit the species based on their abundance within a data matrix. The histograms in Figure 9 corroborate the results of the evenness and diversity measures, i.e., there is a higher evenness in the representation of species in the *muti* shops. Additionally, there are fewer common or abundant species sold by the street traders compared to the *muti* shops and, therefore greater dominance of a few species in samples of the former. Further, street traders sell a smaller range of species that are more prevalent in the market than other species. It is known from research conducted in the market (Williams 2003) that many of these species are currently threatened or have the potential to be so if current harvesting and utilisation trends continue.

An example of intra-sample diversity: healers and non-healer traders from the faraday market

The main use of diversity measures within this paper has been to compare inter- and intra-sample diversity and sampling effort with respect to traders of different ethnic groups selling plants within formal and informal markets. A further example of the application of diversity indices to intra-sample differences is the comparison of plants sold by traders that are traditional healers and traders that are *not* traditional healers within the Faraday market. The data matrix for Faraday was subdivided into 'healers' and 'non-healers' and analysed accordingly. The results are shown in Table 5. The results make it clear as to why it is necessary to describe a sample in terms of more than one index of species richness, evenness and diversity as well as the necessity for a species accumulation curve (Figure 10).

Numerical species richness of the plants sold by healers is lower than non-healers (Table 5) – however, this is a function of the smaller subsample size. The curve for healers lies above the curve for non-healers (Figure 10), and therefore the subsample is comparatively more species rich. Additionally, Margalef's index for species richness indicates that species richness is almost the same for both subsamples, with the fractionally higher value for non-healers being the result of a larger subsample. Shannon's index of diversity is exactly the same for both subsamples, thereby underlining the importance of using additional indices to discriminate between sample diversity where sample sizes differ. Simpson's index and Fisher's α show plant diversity sold by the healers to be higher, and therefore there is lower dominance of species sold by the healers. Additionally, the probability of encountering the same species amongst the non-healer traders is higher – the low evenness values for the non-healer subsample corroborate this evidence. The greater dominance of a few species in the non-healer subsample can be accounted for in the following way: unlike healers, non-healers do not supplement their trading incomes with paid consultations by patients. With the high level of competition in the market, the non-healer traders cannot afford to keep too many species that have intermediate demand and commercial value (wholesale), and they therefore, tend to keep more of the species known to be in demand by customers. The plant knowledge traditional healers have, by contrast, allows more flexibility in the range of plants sold – some of which are added to mixtures and preparations sold to patients.

Conclusions

The true value of diversity measures will be determined by whether or not they are empirically useful (Magurran 1988). In non-ethnobotanical studies there are two main areas in which diversity measures have potential application, namely: conservation management, which concentrates almost exclusively on measures of species richness (underpinned by the idea that species-rich communities are better than species-poor ones); and, environmental monitoring, which makes extensive use of diversity indices and species abundance distributions (where, for example, the adverse affects of pollution will be reflected in a reduction in diversity or a change in the shape of the species abundance distribution) (Magurran 1988).

According to Begossi (1996), diversity measures can be used to evaluate the intensity of resources used by human populations, to allow comparisons among different populations in different environments, and to allow evaluations of sampling effort. Further, Begossi (1996) used diversity indices to help answer the following questions: does the diversity of plant use in an area represent the floristic abundance available?; are the same plants used by most individuals?; and, are there differences in the diversity of plant uses per category of user (e.g., gender/age)?

Diversity indices, in the broad sense of the term, therefore, have a useful role to play in the quantitative analysis of ethnobotanical data. Which measures are 'best' cannot be decided without first knowing why diversity should be of interest (Pielou 1975) or how it can help appraise the availability of plants used and/or traded commercially. The primary goal of this paper was to evaluate the performance of a large variety of indices in relation to samples of different sizes and trader profiles, and to examine the kind of useful information they provided. The primary criterion used for recommending certain measures is the value and economy that can be added to the description of plant availability/use, and therefore the degree of insight that can be acquired into interpreting relative abundances. Second, indices are recommended based on ease of calculation and the extent of use by other researchers so that comparisons may be made with other data sets similarly analysed (e.g., Begossi 1996; Hanazaki et al. 2000). The relative merits and shortcomings of some diversity measures have been previously described in Magurran (1988), and an awareness of their limitations is necessary (Ludwig and Reynolds 1988). Ultimately, the choice of index depends on the requirements of the researcher and the value that the index adds to the quantitative description and understanding of the resources under investigation.

A single index of diversity will most often not be sufficient to describe inter- and intra-sample diversity (Hayek and Buzas 1997). Additionally, to describe a sample only in terms of its diversity/heterogeneity index is to confound the two factors of species richness and evenness (Pielou 1975). It is, therefore, judicious to describe a sample in terms of richness, evenness *and* diversity. To this end, we recommend the use of the following measures: (1) species richness (S or N0); (2) Margalef's index; (3) Shannon index (H'); (4) Simpson's index (both λ and $-\ln \lambda$); (5) Hill's diversity numbers N1, N2 and $N\infty$; (6) Fisher's α; and (7) evenness indices E1 and E5. It is additionally essential that these indices are graphed as diversity accumulation curves so that the performance of an index may be comprehensively evaluated, and the minimum viable sample size can be determined. According to Gotelli and Colwell (2001), comparing richness without reference to a taxon sampling curve is problematic and graphing the results is necessary to detect differences in measured species richness (and diversity) related to the relative abundance shown in the species accumulation or species diversity curves. We recommend the use of *EstimateS* (Colwell 2001) as a basis for calculating the accumulation curves and the input values necessary for computing most of the other indices. Hill's numbers have an additional beneficial use in the delimitation of species that are rare, intermediate and common in abundance within a sample, and this is a crucial first step in prioritising species for monitoring and/or remedial conservation action.

The second objective of the paper was to assess whether the survey sites were adequately sampled, and determine the minimum viable sample size on which a diversity measure should be based for the type of data collected. Rarefaction is a commonly used method for estimating species richness, and can be applied to evaluating sampling effort (Magurran 1988; Begossi 1996; Williams et al. 2000; Gotelli and Colwell 2001). However, the size of the sampling unit should be chosen according to factors other than richness because relative abundance affects the performance of the indices. When evaluating species diversity measures (including richness and evenness) a sample size of at least 20 *muti* shops and 35–40 street traders is necessary for the formal and informal sectors in the Witwatersrand, respectively, (Table 6). The actual number of traders surveyed depends on what aspect of diversity is being measured (Table 6). One reason for the necessity to sample more street traders than *muti* shops is because of the lower mean number of species sold per street trader, therefore, requiring additional sampling effort to increase the number of individuals

Table 5. Comparisons of selected measures of diversity between healer and non-healer traders in the Faraday street market.

Index/Measure	Healer traders (n = 39) N = 1023	Non-healer traders (n = 60) N = 1351
Species richness (S/N0)	281	295
Margalef	40.4	40.8
Shannon (H')	5.25	5.25
Simpson (λ)	0.0062	0.0067
Simpson ($-\ln \lambda$)	5.08	5.00
Fisher's α	127.9	116.4
Hill's N1	190.6	190.6
Hills N2	161.1	148.6
Hill's N∞	40.9	39.7
Evenness E1	0.931	0.923
Evenness E2	0.845	0.779

Table 6. The minimum viable sample size on which a species diversity measure should be based.

Index/Measure	*Muti* shops	Street traders
Species richness	At least 20	At least 40
Diversity indices		
Information theory (e.g., Shannon)	15–25	25–30
Dominance indices (e.g., Simpson)	20+	15–20
Hill's numbers	20+	30+
Fisher's alpha	30	35
Evenness	17	22
Summary: minimum viable sample size	20–30	35–40

recorded for the sampling curves to reach a horizontal asymptote. Overall, however, sampling effort was found to be more than adequate.

The third objective was to compare the species diversity of sites within the formal and informal sector, and thereby appraise plant availability within the ethnomedicinal trade in the region. As a result, inter- and intra-sample similarities and differences in the sale of plants were identified. Numerical species richness was found to be higher for *muti* shops than street traders, despite the smaller sample size (n). This is related to the large number of individuals (N) and ethnospecies sold per shop trader. Most of the diversity accumulation curves for the indices showed a distinction in the availability and relative abundance of plants sold by street traders compared to shop traders. Different trading factors, therefore, operate within the trading sectors to determine the plant diversity for sale. The graphs of the diversity curves are, therefore, essential for interpreting the different mechanisms operating within the different markets. The average degree of uncertainty in predicting the identity of species sold by the traders is higher in the formal sector, therefore, diversity is higher and dominance of a few species is lower. The higher dominance of certain plants sold by the street traders is confirmed by the lower evenness values of the samples. Intra-sample differences in the *muti* shops showed Indian traders to sell a larger diversity of plants compared to the Black traders, and therefore dominance of plants in the latter was higher. In general, all the indices gave different values but consistent rankings for the different sites.

The high diversity of plants sold within both the formal and informal sector in the study area is likely to be related to a number of factors. Cities (like Johannesburg) are more likely to have more culturally diverse populations, drawn in from many rural communities (Cunningham 2001). Diversity of species sold increases with increasing size of the marketing area, and therefore more species are sold in regional markets (such as the Witwatersrand), fewer in central markets and still fewer in minor or local markets (Cunningham 2001). Begossi (1996) suggests that local resistance to Western medicine may result in a greater demand for traditional medicines, thereby increasing the diversity of plants used. At least 12–15 million people are estimated to consult traditional healers in South Africa annually, and urbanisation has not precluded the use of traditional medicine. In one 'township' southwest of Johannesburg (Soweto), there were estimated to be at least 18,000 traditional healers. The Faraday Street market functions primarily as a wholesale market to the traditional healers in townships in the region (Williams 2003). The high diversity of traditional medicines sold in the region is, therefore, indicative of the high demand and the acceptability of traditional healing practices – which to some extent is related to the affordability of primary health care.

Ecologists have long known of species richness, diversity and evenness, and it is only recently that these measures have been applied to the quantitative analysis of ethnobotanical data (e.g., Begossi 1996; Williams et al. 2000). The methods add greater depth to the exploration and understanding of mechanisms and patterns operating within the field of indigenous plant use and trade.

While a quantitative approach to analysing ethnobotanical data might not always be possible, the approach is highly recommended.

References

Begossi A. 1996. Use of ecological methods in ethnobotany. Econ. Bot. 50: 280–289.

Botha J. 2001. Perceptions of Availability and Values of Medicinal Plants Traded on the Western Boundary of the Kruger National Park, South Africa. MSc dissertation, University of the Witwatersrand, Johannesburg.

Botha J., Witkowski E.T.F. and Shackleton C.M. 2001. An inventory of medicinal plants traded on the western boundary of the Kruger National Park, South Africa. Koedoe 44(2): 7–46.

Brower J.E. and Zar J.H. 1977. Field and Laboratory Methods for General Ecology. Wm. C. Brown Publishers, Dubuque, Iowa.

Colwell R.K. and Coddington J.A. 1995. Estimating terrestrial biodiversity through extrapolation. In: Hawksworth D.L. (ed.), Biodiversity Measurement and Estimation. Chapman and Hall, pp. 101–118.

Colwell R.K. 2001. EstimateS: Statistical estimation of species richness and shared species from samples. Version 6, User's guide and application published at: http://viceroy.eeb.uconn.edu/estimates.

Cunningham A.B. 2001. Applied ethnobotany: people, wild plant use and conservation. People and Plants Conservation Manual, Earthscan.

Dauskardt R.P.A. 1990. The changing geography of traditional medicine: urban herbalism on the Witwatersrand, Johannesburg. GeoJournal 22(3): 275–283.

Dauskardt R.P.A. 1991. 'Urban herbalism': the restructuring of informal survival in Johannesburg. In: Preston-Whyte E. and Rogerson C. (eds), South Africa's Informal Economy. Oxford University Press, Cape Town, pp. 87–100.

Gotelli N.J. and Colwell R.K. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol. Lett. 4: 379–391.

Hanazaki N., Tamashiro J.Y., Leitão-Filho H.F. and Begossi A. 2000. Diversity of plant use in two Caiçara communities from the Atlantic Forest coast, Brazil. Biodivers. Conserv. 9: 597–615.

Hayek L.C. and Buzas M.A. 1997. Surveying Natural Populations. Columbia University Press, New York.

Höft M., Barik S.K. and Lykke A.M. 1999. Quantitative ethnobotany: applications of multivariate and statistical analysis in ethnobotany. People and Plants working paper 6. UNESCO, Paris.

Hill M.O. 1973. Diversity and evenness: a unifying notation and its consequences. Ecology 54(2): 427–432.

Johns T., Mhoro E.B., Sanaya P. and Kimanani E.K. 1994. Herbal remedies of the Batemi of Ngorongoro District, Tanzania: a quantitative appraisal. Econ. Bot. 48: 90–95.

Krebs C.J. 1989. Ecological Methodology. Harper and Row Publications, New York.

Lowrey T.K. and Wright S., eds. 1987. The Flora of the Witwatersrand, Volume I, The Monocotyledonae. Wits University Press, Johannesburg.

Ludwig J.A. and Reynolds J.F. 1988. Statistical Ecology – a Primer on Methods and Computing. John Wiley and Sons, Toronto.

Luoga E.J., Witkowski E.T.F. and Balkwill K. 2000. Subsistence use of wood products and shifting cultivation within miombo woodland of eastern Tanzania, with some notes on commercial uses. S. Afr. J. Bot. 66(1): 72–85.

Magurran A. 1988. Ecological Diversity and its Measurement. Princeton University Press, Princeton.

Murali K.S., Shankar U., Shaanker R.U., Ganeshaiah K.N. and Bawa K.S. 1996. Extraction of non-timber forest products in the forests of Biligiri Rangan Hills, India. 2. Impact of NTFP extraction on regeneration, population structure, and species composition. Econ. Bot. 50(3): 252–269.

Phillips O. and Gentry A.H. 1993a. The useful plants of Tambopata, Peru: I. Statistical hypotheses tests with a new quantitative technique. Econ. Bot. 47: 15–32.
Phillips O. and Gentry A.H. 1993b. The useful plants of Tambopata, Peru: II. Additional hypothesis testing in quantitative ethnobotany. Econ. Bot. 47: 33–43.
Pielou E.C. 1975. Ecological Diversity. John Wiley and Sons, New York.
Prance G.T., Balée W., Boom B.M. and Carneiro R.L. 1987. Quantitative ethnobotany and the case for conservation in Amazonia. Conserv. Biol. 1: 296–310.
Williams V.L., Balkwill K. and Witkowski E.T.F. 1997. Muthi traders on the Witwatersrand, South Africa – an urban mosaic. S. Afr. J. Bot. 63(6): 378–381.
Williams V.L., Balkwill K. and Witkowski E.T.F. 2000. Unravelling the commercial market for medicinal plants and plant parts on the Witwatersrand, South Africa. Econ. Bot. 54(3): 310–327.
Williams V.L., Balkwill K. and Witkowski E.T.F. 2001. A lexicon of plants traded in the Witwatersrand umuthi shops, South Africa. Bothalia 31(1): 71–98.
Williams V.L. 1992. An investigation of the herbal medicine or 'muti' trade on Witwatersrand. BSc (Hons) Dissertation, Department of Geography and Environmental Sciences, University of the Witwatersrand, Johannesburg.
Williams V.L. 2003. Hawkers of health: an investigation of the Faraday Street traditional medicine market in Johannesburg, Gauteng. Report to the Gauteng Directorate of Nature Conservation, DACEL, Johannesburg.
Wong J., Thornber K. and Baker N. 2001. Resource assessment of non-wood forest products: experience and biometric principles. NWFP Series 13, FAO, Rome.
Zar J.H. 1984. Biostatistical Analysis. Prentice-Hall, Inc., Englewood Cliffs.