Ecological Sampling Design and Observer Bias: An Example from Toad-Headed Agama (*Phrynocephalus versicolor*) in the Southern Gobi, Mongolia

Kate R. Oddie*,, Togloom Ariunaa* and Davaadorj Enkhnasan***

**Faculty of Biology, National University of Mongolia, Ulaanbaatar 210646, Mongolia.*
***Centre d'Ecologie Functionelle and Evolutive, UPR 9056 CNRS, 1919 Route de Mende, F-34293, Montpellier Cedex 5, France e-mail:gobiology@yahoo.com*

Abstract

Currently Mongolia faces great challenges in ecological research, with ecological studies relatively neglected during Soviet administration. The development of ecological studies requires an appreciation of sampling bias and how this can be avoided. Here we present a case where observer biases are impossible to disentangle from site effects because sample site data are confounded with observer sampling activity. Three volunteers were collected simple body mass and length measures of toad-headed agama, Phrynocephalus versicolor, as part of a wider ecological survey programme in Borzongiin Gobi, South Mongolia. Field data potentially reveal a difference in morphological size in toad-headed agama inhabiting different habitats, but this can only be ascertained through further sampling effort. We give recommendations for future studies.

Key words: *Phrynocephalus versicolor*, ecological field techniques, sampling, observer bias, Borzongiin Gobi, Mongolia.

Introduction

Throughout Soviet dominance until 1990, the biology education system in Mongolia focused on systematics and taxonomy, with very little ecological research. Consequently, although a wealth of systematic knowledge has been gathered for Mongolian species, little is known about dynamic processes occurring in Mongolian ecosystems, population sizes and interactive processes between species (Anonymous, 1998). Studies of ecosystem functioning are developing, but resources for such research are scarce and frequently involve untrained individuals in the collection of data. Furthermore, the scale of the ecological work required over such a large, ecologically diverse and unstudied country often necessitates a large workforce to carry out biological sampling.

Techniques for ecological censusing, either for pure biological study or habitat management and conservation purposes, require unbiased estimates to be collected (e.g. Bibby *et al.*, 1992; Sutherland 1996; Begon *et al.*, 1996), involving for example random sampling and sampling replication (Greenwood, 1996). One possibility for the introduction of bias into ecological data is the use of a number of observers to measure the same factor. Variation in data is generated by observer differences rather than representing true biological variation (e.g. Sauer *et al.*, 1994, Robinson *et al.*, 2000, Spaulding *et al.*, 2000).

We document a case of difficulty encountered in disentangling a potential biological phenomenon from a biased sampling effect during a Mongolian field study of toad-headed agama, *Phrynocephalus versicolor.* Field work was intended to compare ecological parameters at two sites in the Little Gobi Strictly Protected Area A. Species richness, abundance and population distributions in this area are little known and simple inventory work as well as more complex ecological and behavioural studies is urgently needed. Three volunteers from Mongolia and overseas were involved with agama data collection as part of a larger survey programme by international non-governmental organisation (NGO) volunteers. Data were collected as a simple training exercise, but demonstrate well the problems encountered with inadequate experimental sampling design and serve as a lesson for future studies.

Materials and Methods

Field work was carried out in Borzongiin Gobi, South Gobi Province, within Little Gobi Strictly Strictly Protected Area A, during 03-15 June 2001. The environment within the protected area is patchy, with wildlife distributed unevenly between distinct habitats. Two areas of particular interest in a zone of mainly sand and gravel plains are (i) natural springs that provide surface water, and (ii) stands of relict elm, *Ulmus sp.*

Data were collected from 5 km transects walked across one such spring (N 42°29'02", E 105°14'06") and through a stand of 60-70 well-developed *Ulmus* trees (N 42°11'30", E 105°20'40"). The sites are 42 km apart. The spring provides year-round surface water that in spring and summer results in a grassy flood plain used for domestic livestock grazing. In both areas the plants of the genera *Zygophyllum, Stipa* and *Salsola* are ubiquitous (though heavily grazed at the spring). The spring is characterised by the presence of *Chenopodium, Iris* and *Oxytropis.* Prevalent in the shrubby desert and dry river beds of the *Ulmus* were *Kallidium, Nitratia, Cargana* and *Anabasis.* Generally there was a greater density of, and variation in, vegetation at the *Ulmus grove* (Oddie, 2001 unpublished report and own data).

Three volunteers collected mass and body length morphometric data on the toad-headed agama. Field studies on agama were part of a larger survey exercise also comparing vegetation and terrestrial insects between sites. This influenced the logistics of data collection. Eight and ten N-S transects of 5km were completed at the spring and *Ulmus* respectively, beginning at 07:15 and ending between 11:00 and 13:00. The number of agamas caught by different three observers at different sites varied, as demonstrated in Table 1.

Agamas were caught by covering individuals with a cupped hand. They were marked with a small spot of coloured nail varnish on their dorsal surface between their rear legs in order to prevent pseudoreplication by distinguishing those already captured (recaptured individuals are not included in analyses).

Agamas were measured to the nearest point from nose to tail by using a metal ruler attached to a wooden platform with one raised end at right angles to the surface, an arrangement commonly used for measuring bird wing lengths. With the nose just touching the perpendicalar section, the body was stretched out flat with the tail being pulled straight and flat next to the ruler before the measurement being read at the tip of the tail. Masses were recorded by suspending agamas in paper cones from pesola spring balances, accurate to 0.1g.

Table 1. Number of agamas caught by three different observers at two sites in Borzongiin Gobi (date is day of June)

Date	Site	N agamas	Observer
3	Spring	21	1
3	Spring	21	2
4	Spring	11	1
4	Spring	23	2
5	Spring	15	2
5	Spring	28	3
6	Spring	6	2
6	Spring	18	3
8	*Ulmus*	6	1
8	*Ulmus*	11	3
9	*Ulmus*	6	1
9	*Ulmus*	11	3
11	*Ulmus*	40	1
11	*Ulmus*	2	2
12	*Ulmus*	11	1
12	*Ulmus*	3	2
13	*Ulmus*	3	3
14	*Ulmus*	6	3

Results

The total number of agamas caught was 242, and mass and length were recorded in 232 and 235 cases respectively. Mass varied between 0.9 and 8.6g, with a mean of 3.54 (±1.47) g. Nose to tail lengths varied between 55 and 50mm, with a mean of 95.60 (±14.52) mm. There was no difference in the overall time of capture between the two sites (t=1.67, p=0.095, d.f.=237).

Simple Wilcoxon tests surprisingly revealed strong differences in body masses and lengths between sites (mass: $Z_{1,232}$=-7.16, p<0.0001, length: $Z_{1,235}$=-6.07, p<0.0001; Fig. 1), with agamas caught at the spring larger and heavier than those in the *Ulmus*.

However, observers also accounted for significant variation in agama masses and lengths (Kruskal-Wallis test, mass: $\chi^2_{2,232}$=10.65, p=0.005, length: $\chi^2_{2,232}$=9.05, p<0.011; Fig. 1). Since these factors are not independent (observers did not catch the same number of agamas at each of the two sites) it is impossible to tease apart variation due to site and variation due to observer.

An ANOVA including both site and observer as factors revealed that even when controlling for differences between observers, both lengths and masses differed between sites (mass: $F_{1,228}$=6.94, p<0.0001, length: $F_{1,231}$=34.42, p<0.0001).

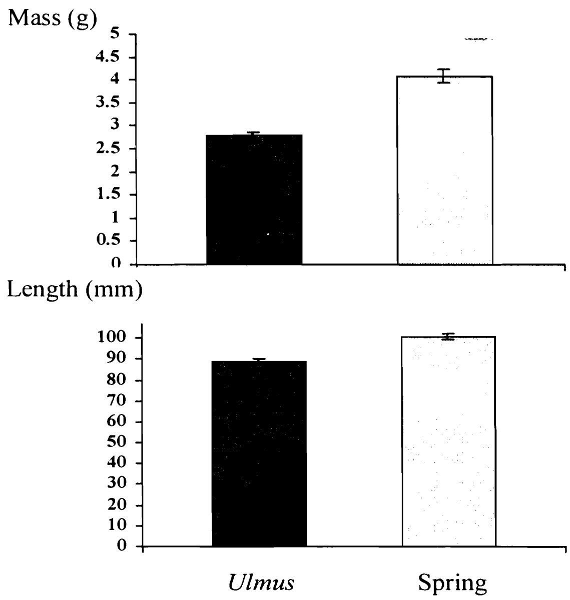

Fig. 1. Differences in mean body (a) mass and (b) nose-tail length (± s.e.) in *Phrynosephalus versicolor* agamas caught at two different sites in Borzongiin Gobi, three observers pooled

However, we are cautious in interpreting whether this difference reflects a true biological phenomenon, given the highly significant differences found between observers (Fig. 2).

Testing the interaction term site observer reveals that observer effects on mass were constant across sites ($F_{1,226}$=1.69, p=0.187) but that the observer had different effects on length at different sites ($F_{1,229}$=3.24, p=0.041).

Plotting the lengths and masses of agamas caught according to observer (Fig. 2) shows that observer 1 measured consistently smaller than the

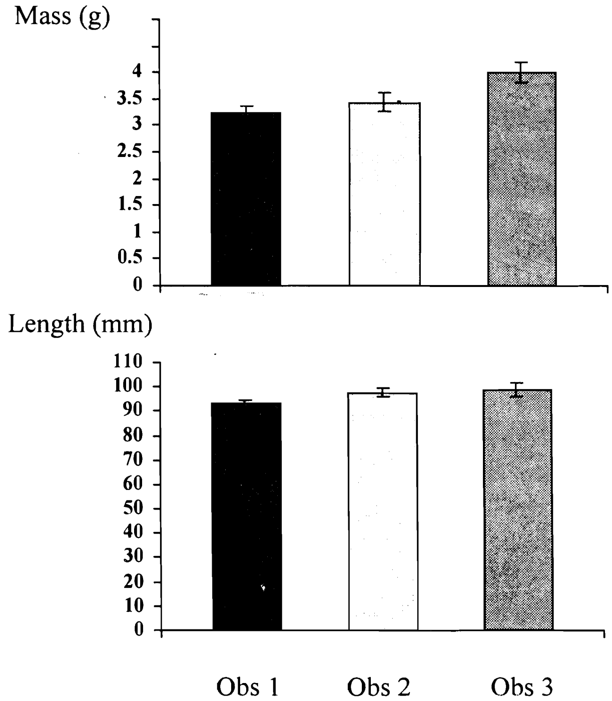

Fig. 2. Observer differences in mean body (a) mass and (b) nose-tail length (± s.e.) in *Phrynocephalus versicolor* agamas caught in Borzongiin Gobi, both sites pooled

other observers. Hence and over-representation of agamas caught by observer 1 at the *Ulmus*, or over-representation of agamas caught by observer 3 at the spring, may account for the differences in measurements between sites. We calculated a contingency table of the number of agamas caught at each of the sites by each observer (Table 2). We then tested whether there was an association between group and site. We found that indeed the number of agamas caught by each group differs highly significantly between sites (χ^2 =58.40, p=0.0001), i.e. sites were sampled unevenly by observers. Hence, agama body size differences may be due to a high proportion of measurements made

Table 2. Contingency table of number of agamas caught by three different observers at two different study sites in Borzongiin Gobi

Observer	*Ulmus*	*Spring*	Total (by observers)
1	63 (26.0)	32 (13.2)	95
2	5 (2.1)	65 (26.9)	70
3	31 (12.8)	46 (19.0)	77
Total (by sites)	99	143	

by observer 1 at the *Ulmus* compared to a relatively numerous measurements made by observers 2 and 3 at the spring (Table 2 and Fig. 3).

It is likely that consistent differences in measures between observers bias these data. Observer differences in nose-tail lengths may be expected as

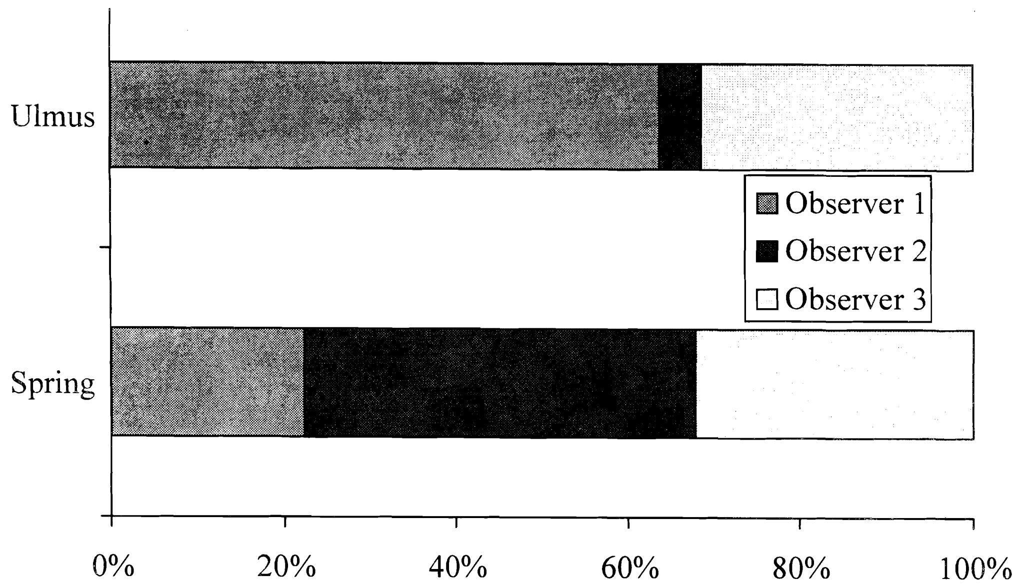

Fig. 3. Proportion of agama captures made by each volunteer per site total

individuals become accustomed to holding agamas in their own particular way, but using a spring balance should produce consistent results across observers. Thus we may expect some error in length measurements, whereas weight measures are generally thought to be more reliable. In this study we find greater variation in mass than in length measures. This could be considered evidence to support body length differences between the two sites. However, it is also true that inexperienced users may have inadequately sheltered spring balances from the frequent Gobi winds, resulting in some variation. Furthermore, it may be that, for at least part of the study, spring balances may not have been correctly zeroed.

We found that the number of agamas caught at each site was not consistent across observers. This represents the second major problem of the study: uneven sampling. In any ecological study it is important to remove as many sources of bias as possible (e.g. Begon *et al.*, 1996; Sutherland, 1999, 2000), or in a study such as this comparing sites it may not be critical if counts at each site are consistently biased. Here the fact that measures were not distributed evenly across sites by observers meant that measurement biases were not evenly represented across sites. In this case, biased effort across sites could not be avoided because of

Discussion

Simple morphometric data collected for the toad-headed agama in Borzongiin Gobi revealed individuals caught at a spring to be both larger and heavier than those caught at an Ulmus grove. However, whether this difference represents a true biological phenomenon or is due to measurement biases between multiple observers remain to be determined, since with current data we are unable to discern if this is due to sampling and observer biases. Three different individuals measured agamas with significant variation between observers, and sites were sampled unevenly by observers.

Given the short distance of only 42km between sites, such large morphological differences as recorded here seem unlikely. It may be that a greater abundance of insect prey at the spring influenced agama measures, resulting in heavier individuals caught at the spring. Differential feeding activity is likely to cause variation in body mass within a species, yet here we also record a surprisingly large variation in body length also. We speculate that differences between sites may be due to two problems involved in experimental design: uneven sampling and use of inexperienced personnel.

logistical constraints and nature of the work. Agama surveys were carried out within teams of five or six young volunteers also collecting data on plants and insects, and groups had to be maintained for safety reasons. Other sources of error such as daily weather differences (that affect agama activity and likelihood of capture) were minimised by carrying out surveys during the same time period each day. Furthermore, sampling was repeated over a number of days at each site to reduce error, rather than relying on data from just one or two sampling days.

These data call into question the usefulness of volunteers to collect data. The use of untrained personnel limits biological studies in both the complexity of data that can be collected and the reliability, hence its usefulness. In this study volunteers received a two-day training period and required reminding of methods and motivating throughout. Using non-trained personnel to collect data may be tempting given that they can cover a wide survey area in a short time, however, variation in observer effort and ability justifies great caution in the use of non-trained personnel (Kendall et al., 1996; Robinson et al., 2000). Consequently, their use in serious biological field studies, especially those upon which conservation decisions rest, should be questioned. If volunteers are to be used, we recommend extensive training, pilot studies and that observer differences are quantified (for example here a sample of at least 20 agamas are measured by all three observers to reveal inconsistencies in measurements) and caution in results interpretation. Inter-observer variation should always be investigated. In this case the work fulfilled concurrent social and educational project aims as observers were involved in a youth expedition to promote a sense of awareness of the natural environment, encourage interest in environmental processes and cross-cultural exchange. Hence, although biological findings may be questioned, the project upheld other aims. Future studies in Mongolia using untrained students or volunteers must determine at the outset the motives for their work and design adequate training, sampling and monitoring activities accordingly.

For the agama, and indeed for other species in Borzongiin Gobi, further investigations are encouraged. These could be coupled with, for example, behavioural observations to record prey items consumed at each site and complementary studies of prey species once identified. Inadequacies in the current agama data mean we are unable to say with certainty that body size differences are caused by biological variation between sites, neither can we conclude they are caused by observers. In future, measurements made by one observer at both sites are required. Furthermore, any observer collecting data should measure at least 20 individual agama two times, in order to calculate a repeatability estimate for his/her own measures (see Lessells & Boag, 1987). If observers cannot measure reliably then the validity of any data are questioned; here further training and practice may be advocated.

Volunteers are useful to collect pilot data, perhaps maybe more so in collecting count data rather than measures. However, more experienced scientific personnel with an understanding of rigorous scientific sampling methods are required to produce reliable data for Mongolian populations. This is especially true for those upon which conservation decisions rest.

Acknowledgements

We thank volunteers and staff of Raleigh International and Blandine Doligez for field work, and the National University of Mongolia and National Parks Office in Dalanzadgad city for support. We especially thank herding families within the study area for their endless help. Work in the Little Gobi SPA was permitted by the Protected Areas Bureau of the Ministry of Nature and Environment of Mongolia. The senior author was supported by a grant from The Leverhulme Trust, London, with start-up sponsorship from the Association for the Study of Animal Behaviour, British Ornithologists Union, Dan Lipscombe Memorial Trust, Edinburgh Davies Trust, Percy Sladen Memorial Trust, Sokkia Communications and the Winston Churchill Memorial Trust.

References

Anonymous. 1998. Biological Diversity in Mongolia. First National Report. Global Environmental Facility, United Nations Development Programme, Ministry for Nature and the Environment of Mongolia. Ulaanbaatar.

Begon M., Harper J.L. & Townsend C.R. 1996. Ecology. Third Edition. Oxford: Blackwell Science.

Bibby C.J., Burgess N.D. & Hill D.A. 1992. Bird census techniques. Academic Press, San Diego, USA.

Greenwood J.J.D. 1996. Basic Techniques. In: Sutherland, W.J. (ed) Ecological Census Techniques. Cambridge University Press, Cambridge UK, pp 11-110.

Kendall W.L., Peterjohn B.G. & Sauer J.R. 1996. First-time observer effects in the North American breeding bird survey. Auk, 113: 823-829.

Lessells C.M. & Boag P.T. 1987. Unrepeatable repeatabilities: A common mistake. Auk, 104: 116-121.

Oddie K.R. 2001. Simple biological surveys at a spring and Populus grove in the Gobi Strictly Protected Area A. A Report for the Mongolian Ministry of Nature and Environment National Parks Authority.

Robinson D.A.Jr., Jensen W.E. & Applegate R.D. 2000. Observer effect on a rural mail carrier survey population index. *Wildlife Soc. Bull.* 28: 330-332.

Sauer J.R., Peterjohn B.G. & Link W.A. 1994. Observer differences in the North American breeding bird survey. *Auk,* 111: 50-62.

Spaulding R., Krausman P.R. & Ballard W.B. 2000. Observer bias and analysis of gray wolf diets from scats. *Wildlife Soc. Bull.,* 28: 947-950.

Sutherland W.J. 1996. *Ecological Census Techniques*. Cambridge University Press, Cambridge UK.

Sutherland W.J. 2000. *The Conservation Handbook. Research, Management and Policy.* Blackwell Science, Oxford UK.

(Accepted: December 2002)