CrossMark

ORIGINAL PAPER

How useful are volunteers for visual biodiversity surveys? An evaluation of skill level and group size during a conservation expedition

Andrea Albergoni[1,4,5] · Ian Bride[1] · Charles T. Scialfa[2] ·
Merlijn Jocque[3,4] · Steve Green[1,4]

Abstract The ability of volunteers to undertake different tasks and accurately collect data is critical for the success of many conservation projects. In this study, a simulated herpetofauna visual encounter survey was used to compare the detection and distance estimation accuracy of volunteers and more experienced observers. Experience had a positive effect on individual detection accuracy. However, lower detection performance of less experienced volunteers was not found in the group data, with larger groups being more successful overall, suggesting that working in groups facilitates detection accuracy of those with less experience. This study supports the idea that by optimizing survey protocols according to the available resources (time and volunteer numbers), the sampling efficiency of monitoring programs can be improved and that non-expert volunteers can provide valuable contributions to visual encounter-based biodiversity surveys. Recommendations are made for the improvement of survey methodology involving non-expert volunteers.

Keywords Volunteer data · Conservation voluntourism · Observer performance · Visual survey · Sampling efficiency

✉ Andrea Albergoni
albergoni.andrea@gmail.com

[1] Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent, Canterbury CT2 7NR, UK

[2] Department of Psychology, University of Calgary, Calgary, AB T2N 1N4, Canada

[3] Royal Belgian Institute of Natural Sciences, ATECO, Vautierstraat 29, 1000 Brussels, Belgium

[4] Operation Wallacea, Wallace House, Old Bolingbroke, Spilsby, Lincolnshire PE23 4EX, UK

[5] Rough Common Road 127, Canterbury CT2 9BS, UK

⌂ Springer

Introduction

Scientists have engaged volunteers increasingly to assist and support their research (Silvertown 2009). Data collected by volunteers often contribute significantly to research projects, especially when guided by experienced scientists (Foster-Smith and Evans 2003). This 'citizen-science' model has also been applied within the tourism industry to generate what has been termed 'conservation voluntourism', 'scientific tourism' or, relatedly, the 'conservation holiday' (Brown and Morrison 2003). Volunteer data collection schemes have been successfully adopted by organizations such as Earthwatch and The School for Field Studies. This form of volunteer program has proven to be a valuable option for conservation research from a financial point of view (Brightsmith 2008), but is also beneficial with regards to the completion of ecological studies (Holt et al. 2013).

The accuracy and consistency of data collected by volunteers is a critical aspect of these projects, as data are often used to support scientific publications and management planning decisions. Many components of scientific research can be learned relatively quickly and volunteers, if sufficiently trained, can gather high quality data (Foster-Smith and Evans 2003; Newman et al. 2003). However for some tasks, the learning is more protracted, with expertise slow to accumulate and the data collected by novice volunteers has sometimes been questioned (Cohn 2008; Léopold et al. 2009).

A large part of volunteer contributions to research in the tropics entails reinforcing biodiversity surveys of understudied natural regions. To estimate the abundance of many terrestrial organisms including herpetofauna in tropical forests, visual encounter surveying is one of the most common and efficient techniques (Doan 2003). The distance sampling method is widely applied to estimate population size and/or density of targeted species (Fewster et al. 2009). This entails employing a 'line transect' method, in which observers walk along transects estimating or measuring perpendicular distance of each detected animal from the center of the transect (Cassey and McArdle 1999). Assessment of population abundance and density are closely related to estimations of the distance from the transect (Fewster et al. 2009). As such, inaccurate detection or estimations of distance are likely to result in biased assessments of population size and/or density. Furthermore, modeling the probability of detection is an important step in the analysis of distance sampling data (Thomas et al. 2010).

Detection issues have historically been ignored (Stauffer et al. 2002), but have been acknowledged more recently as important factors in survey methodologies and statistical models (MacKenzie and Kendall 2002; de Solla et al. 2005; Lind et al. 2005). Although some studies have highlighted areas of unreliability in data collected by volunteers, few objective comparisons between novice volunteers and experts have been conducted in the field (Fitzpatrick et al. 2009). There are many observer-related factors that can affect the accuracy of data collection in biodiversity surveys. Understandably, some researchers argue that the experience of the person collecting data has an impact on the ability to detect target species (MacKenzie and Kendall 2002; Fitzpatrick et al. 2009), but data accuracy can also be affected by an individual's abilities and characteristics, irrespective of skill level (Newman et al. 2003; Pipino et al. 2002; Schmitt and Sullivan 1996). If novice volunteers' detection rates are low, this introduces an undocumented source of variation and bias (Fitzpatrick et al. 2009). Thus, the relationship between observer experience and detection in field research requires further investigation (McCarthy et al. 2013).

Detection can also be affected by characteristics of the survey protocol, including duration and number of participants (Gooch et al. 2006; Schmeller et al. 2009). It has been suggested that a larger volunteer sampling effort, which increases with group size, could

counterbalance measurement errors in the data collected (Hochachka et al. 2000). But, surprisingly, the relationship between survey duration and detection efficiency has received little attention (Pierce and Gutzwiller 2004). Similarly, being able to determine the optimum group size has important implications for the way monitoring protocols are designed, especially when human and temporal resources are limited (Ryan et al. 2002). The quality of data collected by volunteers is, in fact, more likely to be affected by survey protocols and design than by volunteer ability per se (Schmeller et al. 2009). Understanding the optimal sampling effort required for specific studies could reduce the likelihood of missing specimens during data collection (de Solla et al. 2005). It will also have important implications for the design and organization of voluntourism projects.

Against this background, we examined the efficiency of volunteers to detect amphibians and reptiles in a tropical forest. We used an experimental design with imitation animals along a transect in the Honduran cloud forest during a conservation 'voluntourism' expedition and compared detection and distance estimation accuracy of volunteers and more experienced observers in different group sizes. This study had two major goals; (1) to evaluate the importance of skill level (along a gradient from high school students to trained scientists and local guides) on biodiversity data collection and (2) to examine the effects of group size on the accuracy of survey data.

Materials and methods

Study site

Cusuco National Park (CNP) is a 23,400-hectare protected area in the Merendón mountain range in northwestern Honduras and consists mostly of upper montane forest. The park supports considerable biodiversity. Cusuco NP is part of the Mesoamerican biodiversity hotspot (Olivet and Asquith 2004) and designated as a key Biodiversity Area by the IUCN. More specific towards amphibians and reptiles it is recognized by Alliance for Zero Extinction for the critical habitat it provides to six endemic Honduran amphibian species
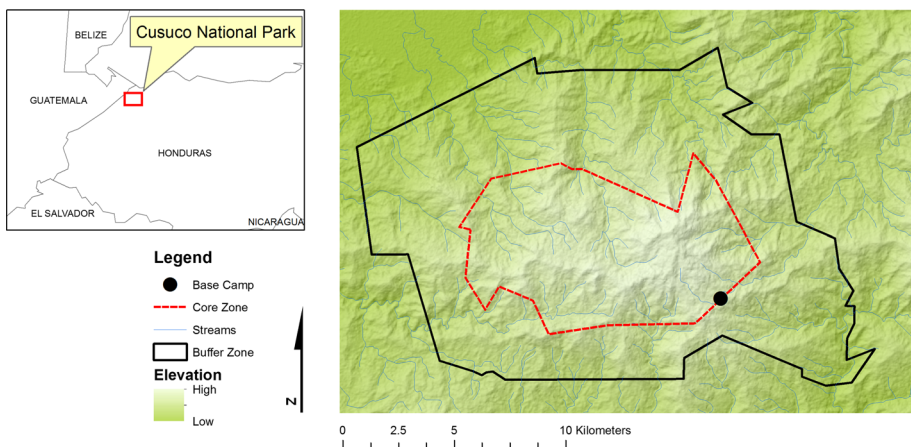

Fig. 1 Map of Cusuco National Park and wider surroundings. The experimental transects were located close to 'Base Camp' (N 15°29′47″, W 88°12′43″) within the core zone of the park

and is listed as the 25th most irreplaceable area for threatened amphibians worldwide (le Saout et al. 2013). The transect was situated in the southeastern core area of the park at 1550 m in proximity to the base camp (see Fig. 1). Data were collected between June 13th and July 21st, 2014.

Participants

Each year Operation Wallacea, a UK-based, volunteer-driven conservation organisation, monitors the biodiversity in CNP along a fixed set of transects emanating from different

Fig. 2 *Upper*—Plastic reproduction herpetofauna used for experiment. *Middle*—Model placed along transect. *Lower*— View of transect

camps. Surveys are led by professional scientists joined by volunteers, including high school students and university students. High school students, who generally only spend a week in the Park accompanied by their teachers, undertake their own programs as school expeditions and go through skills training, academic lectures and practicals to demonstrate the differing types of surveys being undertaken. University students, who spend between 2 and 8 weeks in the forest depending on their objectives, join the program to strengthen their resumé, gain course credit, or collect data for a dissertation or thesis. High school students, university students, professional scientists specialised in different taxa (therefore considered experienced in collecting ecological data using various monitoring techniques) and local guides were all invited to survey the experimental transect.

A total of 280 people were involved in the study; 238 student volunteers (181 high school students and 57 university students), 30 members of staff (mostly scientists and university academics specializing in different taxa), 6 experienced herpetologists and 6 local guides. Participants ranged in age from 16 to 43 years, however the great majority was not older than 25 years. Operation Wallacea provided ethics approval to work with student volunteers, and all participants were made aware of purpose of the study before undertaking the experiment.

Experiment design

The experiment consisted of a simulation of a visual encounter survey for herpetofauna, in which plastic reproductions of amphibians and reptiles were placed at varying distances along a path in the forest. Participants were asked to walk the transects with the goal of detecting the models and estimating their positions (perpendicular distance from transect and distance along the transect line). Two transects (herein labeled A and B), each 200 m in length, were placed along a footpath in an otherwise unused area of the forest. Along each transect, 15 plastic reproductions of frogs, lizards and snakes were securely placed using "zip ties" in as natural a position as possible on adjacent vegetation. The models were placed in three height categories. On both transects, five models were placed at 'ground' level (0–0.5 m above the ground), five at a 'middle' level (0.5–2 m above ground) and five at a 'top' level (over 2 m above the ground). Within each of these three categories the models were placed at 0, 1, 2, 3 and 4 m from the imaginary centre line of the transect. Thus, on each transect, there was one model placed randomly at each combination of height and distance. An equal number of models were used for each taxa and the models were roughly equally distributed on both sides of the transect mid-line. The models were of many different colors, did not resemble any particular species present in the park, and measured 5 cm (frog), 15 cm (lizard), or 20 cm (snake) in length (see Fig. 2). Each transect covered an area of 1600 m^2, giving a density of 1 model per 106.7 m^2. This density is not extremely low or high with respect to field expectations (Dodd and Dorasio 2004).

Before undertaking the experiment, examples of the models were shown to participants for a brief period of time (less than a minute) and instructions on how to complete the recording sheet were provided. Participants were asked to walk the 200 m transect at their chosen pace and try to detect the models, recording the taxonomic group (e.g., snake) and the color. They were also asked to estimate the distance of the models from the center of and along the transect. To facilitate the estimation of distance along the transect, the path was marked every 25 m with pink flagging tape. Local guides were asked only to detect species and estimate the perpendicular distances from the transect. As the models did not resemble any particular species present in the park, participants were not required to

identify any of the models. Data were collected between 9 a.m. and 12 a.m. or between 1 p.m. and 4 p.m. in order to limit variation in ambient illumination.

Participants were divided into five different groups depending on an assumed level of expertise based on their familiarity with the survey technique and the environment they were working in. Listed in increasing order of expertise, the groups were; high school students, university students, members of staff, experienced staff (herpetologists) and local guides. Local guides, given the fact that were born in or near the park and regularly engaged in daily activities in the environment, were considered (albeit with some uncertainty) the most experienced group in regards to detection ability. To establish the effect of group size, participants from each skill group were divided into subgroups of 1, 2, 4 or 8 observers. The numbers in each combination of group size and skill level are shown in Table 1. When tested individually, individual responses were collected. However, when tested in groups, there was only one response per group. (See the Data Collection and Analysis section below). All groups completed the experiment on transect A first and were then randomly reshuffled to create new groups and repeat the experiment on transect B.

Data collection and analysis

For every "survey" we collected the following data: survey time (minutes), group size (1, 2, 4, 8), skill level (high school, university, staff, scientists, guides), each observation of a species, height (low, medium, high), distance from the transect (0, 1, 2, 3 or 4 m), estimated distance along transect and perpendicular distance from the center of the transect.

The central dependent variables analyzed were the number of models detected and the average accuracy of estimated distance along and perpendicular to the transect. Distance accuracy was operationalized as the mean absolute value of the difference between estimated and objective distance. There were two indicators, one for distance along transect and one for perpendicular distance from the center of the transect.

Given this description of the data, one might expect them to be analyzed using an omnibus univariate general linear model in which the effects of group size, skill level, distance and height are examined simultaneously, allowing for main effects and interactions. This was not feasible for several reasons. First, all participants did not provide all data. For example, the guides did not estimate distance along the transect for detected targets. Of greater importance, the number of individuals providing data for each "cell" of the design varied greatly because there were far more volunteers than staff, herpetologists and guides. This very common reality in fieldwork, in which many cells contain no data (e.g., guides working in larger groups) created a non-orthogonal design. There is controversy about the use of "ignoring" and "allowing" tests in such cases (Maxwell and Delaney 2003). As such, although variants on the general linear model were employed for

Table 1 Experiment replicates for the different subgroups

Group skill Group size	High school	University	Staff	Herpetologist	Guide	Tot
1	10	5	24	6	6	51
2	21	12	9			42
4	17	9				26
8	20	9				29

all analyses, there were somewhat separate analyses used to examine the main effects of group size and skill level.

Tests of the main effects of skill and group size were conducted using between-subjects Analyses of Variance (ANOVA) and Analyses of Covariance (ANCOVA), the latter to control for differences in search duration. To test the effect of skill level, performance was examined with one analysis that ignored group size and included all skill levels (hereafter named the 'group dataset', because only one response was recorded for each group) and a second analysis that included only the data from those working as individuals, which included all skill levels (hereafter referred to as the 'individual dataset'). All analyses were performed in SPPSv.22 (IBM 2013).

Results

The experiment was repeated 148 times, with a total of 844 targets detected. On average, observers detected 38 % of the models and ranged from 0 to 75 %, comparable to other work (Foster-Smith and Evans 2003). The majority of models detected were snakes (41 %), followed by frogs (33 %) and lizards (26 %). Not suprisingly, the largest models were detected with greatest frequency. Models were detected with greater likelihood in the middle level (43 %), followed by ground (29 %) and top (28 %).

Survey time

No significant difference was found in the mean number of models detected or mean time spent walking each of the two transects (independent t test $p > 0.05$) and, therefore, data for the two transects were combined.

In the group data, the mean time spent to walk a single transect was 29.6 min. It differed significantly among the five skill groups ($F = 4.81$, $p = 0.001$). Post hoc tests found the herpetologists (mean$_{\text{time spent}}$ = 37.83 min.) spent statistically more time than both high school (mean$_{\text{time spent}}$ = 28.48 min.) and university students (mean$_{\text{time spent}}$ = 28.03 min.). In the individual data, time spent walking the transect was related to group skills ($F = 2.63$, $p = 0.047$). However the post hoc tests revealed that the only significant difference was between the performance of the herpetologists and high school students ($p = 0.026$).

When analyzing the group data, the time taken to walk the transect was also correlated with group size ($r = -0.208$, $p = 0.011$), with larger groups spending less time to complete the task.

Detection accuracy

Effects of skill level

Time spent on the transect had a moderate positive correlation with the number of models detected for both the group data ($r = 0.199$, $p = 0.015$), and individual data ($r = 0.389$, $p = 0.04$). Consequently, to examine detection accuracy as a function of skill level, a between-subjects Analysis of Covariance (ANCOVA) was conducted, using time spent as the covariate.
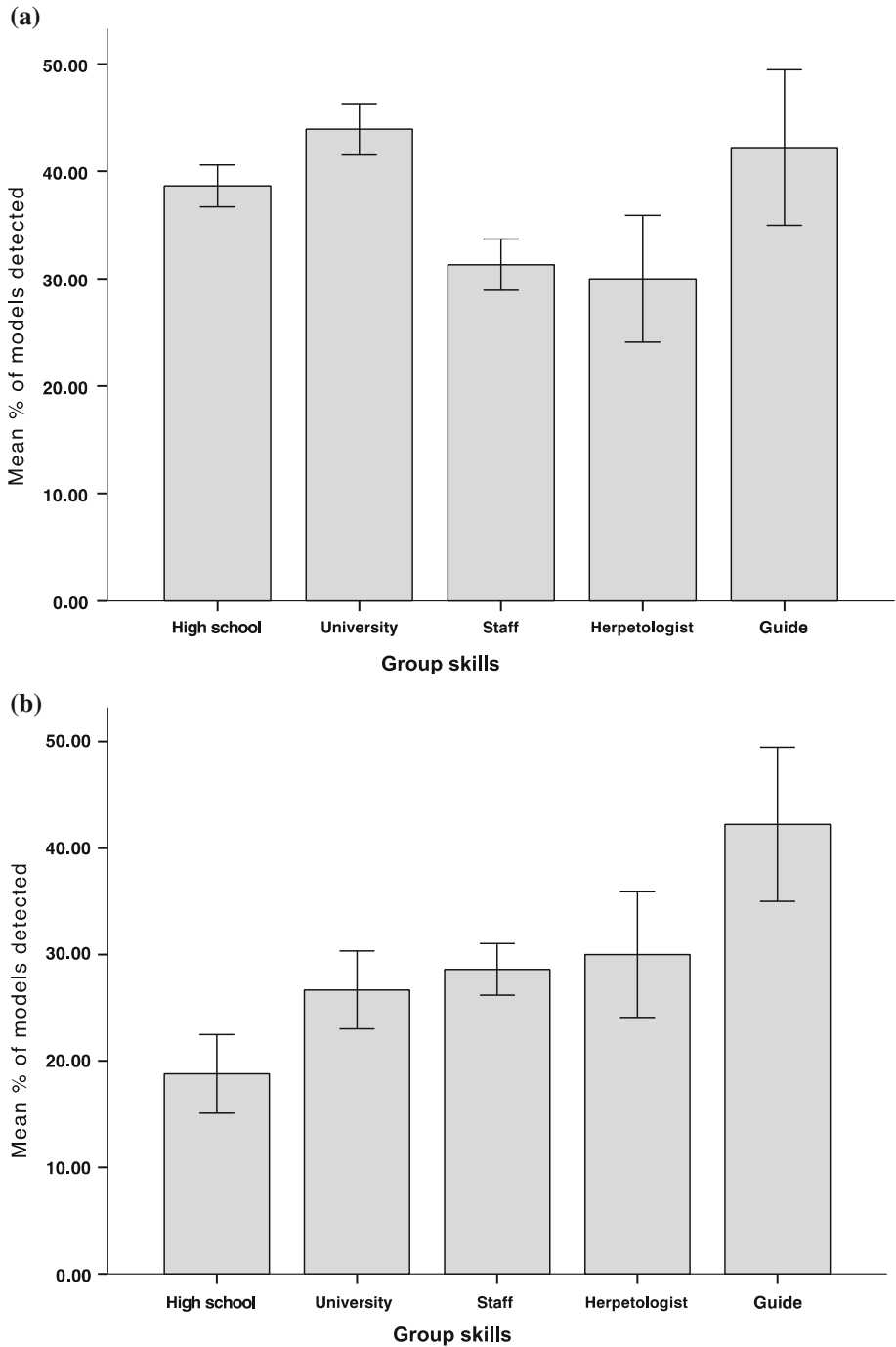
(a)

(b)

Fig. 3 Mean percentage of models detected for the different group skills in **a** group dataset and **b** individual dataset. *Error bars* are 1 *SE*

In the group data, the ANCOVA revealed that time had a significant effect on detection ($F = 14.5$, $p < 0.001$), but skill level still had a significant effect after controlling for time spent searching ($F = 5.74$, $p < 0.001$). Counter to expectations, it does not appear that detection accuracy improves with presumed experience (See Fig. 3, upper panel). In fact, post hoc tests revealed the only statistically different groups to be the university students ($mean_{detected\ models} = 43.9$ %) and the staff members ($mean_{detected\ models} = 31.3$ %), and in this case, it was those with less experience who performed best. It must be borne in mind, however, that the less experienced students were more likely to work in groups and, as will be shown below, this has a facilitative effect on search performance.

A different picture emerges within the individual data (See Fig. 3, lower panel). Time spent on the survey explained some of the variance in performance ($F = 5.97$, $p = 0.018$), but the ANCOVA showed that skill level still had a significant facilitative effect on detection ($F = 2.59$, $p = 0.049$). It is clear that those with more experience were more likely to detect targets, although there are only trivial differences among university students, staff and herpetologists.

Effects of group size

The effect of group size on detectability was examined with a between-subjects ANCOVA, using time as the covariate. After controlling for differences in search time, group size still had a significant effect on detection accuracy ($F = 34.8$, $p < 0.001$). Examination of Fig. 4 shows that larger groups were able to detect more targets. Post hoc tests revealed a significant difference in the mean percentage of models detected between all group sizes, with the exception of groups sized four and eight, which were statistically similar. The same test conducted only on data collected from high school and university students (the most densely populated cells in the design), showed similar results ($F = 31.5$, $p < 0.001$), with only groups of four ($mean_{detected\ models} = 45.6$ %) and groups of eight ($mean_{detected\ models} = 51.7$ %) statistically indistinguishable from each other.

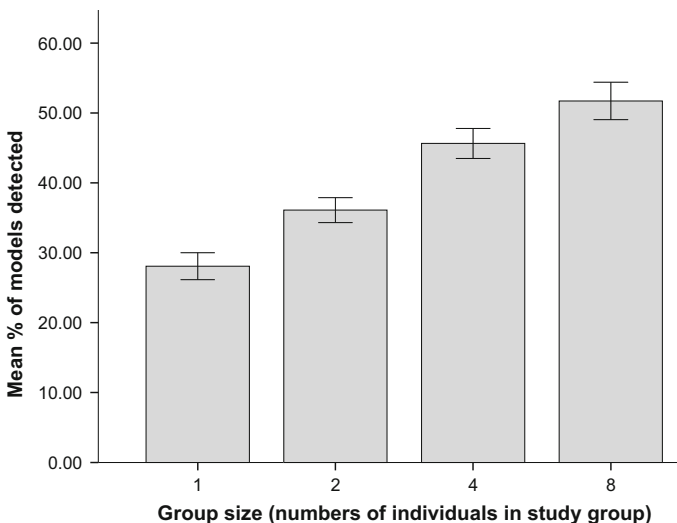Fig. 4 Mean percentage of models detected for the different group sizes. *Error bars* are 1 *SE*
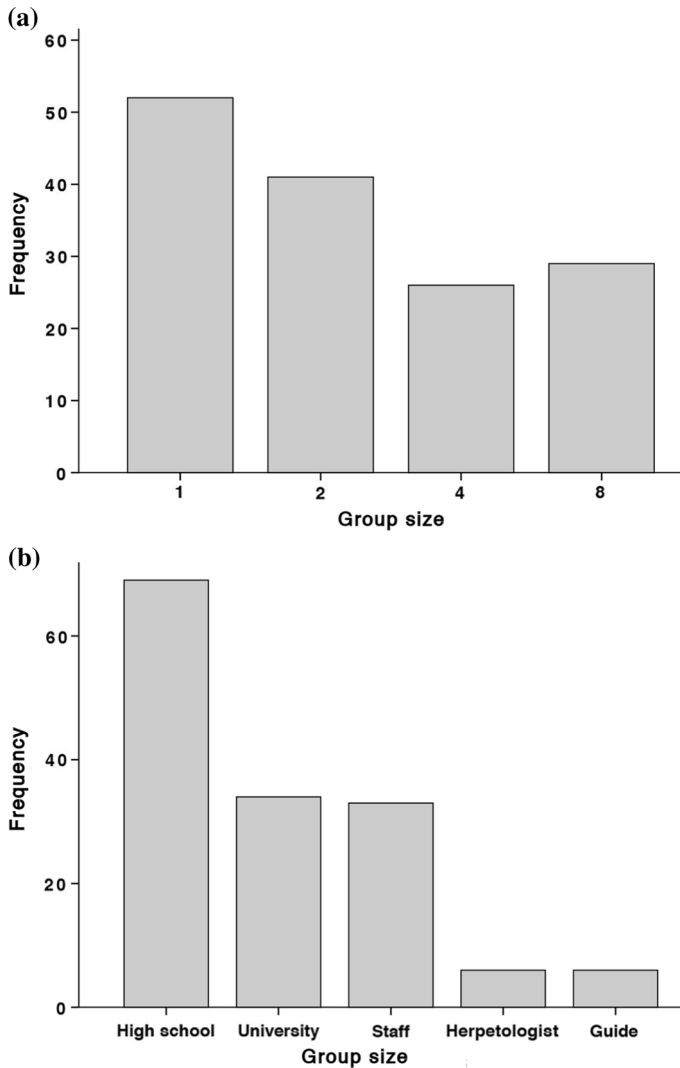
(a)

(b)

Fig. 5 Frequency of the number of trials for different **a** group size and **b** group skills

Thus, the poorer performance of less experienced people seen in Fig. 3 (lower panel) appears to be mitigated (Fig. 3, upper panel) by searching in larger groups. In fact, the number of trials varied significantly between skill groups and group size (see Fig. 5), because there were many more student volunteers than staff available to participate.

Distance estimation

Independent t tests examined differences in estimation accuracy between the two transects. No difference ($p = 0.21$) was found for the accuracy of perpendicular distance estimations between transect A ($mean_{error} = 0.45$ m) and B ($mean_{error} = 0.52$ m) There was, however, a difference ($p = 0.02$) in the accuracy of estimations for distance along the transect, with

a slight improvement in transect B (mean$_{error}$ = 1.26 m) compared to transect A (mean$_{error}$ = 1.66 m). This difference may well be due to practice, because transect order was fixed and transect B always surveyed last.

Because time spent was not correlated with either measure of distance accuracy (perpendicular distance $r = 0.09$, $p = 0.27$; distance along transect $r = 0.08$, $p = 0.35$), ANOVAs were used to assess whether there was a skill or group size effect on estimating the position of models (with the exception of guides, who were tested only for perpendicular distance estimation). The ANOVAs were run on the group data set, on data collected by individuals only and on data collected by school and university volunteers of all group sizes (as they represented all group skills). Neither group size nor group skill had a significant effect on estimation accuracy ($ps > 0.16$). Results for the group dataset are shown in Fig. 6.

Discussion

This study found skill level to have a positive effect on detection, corroborating the view that the number of detected targets in a survey is associated with expertise, particularly in the case of low-density populations (Shirose et al. 1997; Fitzpatrick et al. 2009). However, this effect was observed only when analyzing the individual data. It should be borne in mind that skill, as operationalized here, is a combination of familiarity with the task and the context. This can be seen in the observation that local guides searching as individuals had the best detection rates.

Our study also illustrates how detection performance of relatively untrained volunteers can be augmented by using larger groups. Thus, these data corroborate the notion, already discussed by Freilich and LaRue Jr (1998), that inexperienced volunteers can perform straightforward tasks, such as a visual detection survey, as competently as more experienced observersm be it in larger numbers.

In contrast, the ability to accurately estimate distance was no different between experienced and inexperienced observers, suggesting that experience alone does not ensure greater accuracy in the survey estimations. This finding is at odds with other evaluations, (Shirose et al. 1997; Alldredge et al. 2007). It should be emphasized that the presence of longitudinal distance markers greatly simplified the estimation of distance along the transects and that distance estimation at these distances is known to be quite accurate, even among untrained observers (Wiest and Bell 1985). There may well be group and/or skill differences in more demanding estimation tasks. Additionally, local guides were not asked to estimate the distance of models along the transect.

Elements of the survey protocol were also examined to understand which ones might have an effect on detection and accuracy of estimations. A significant effect of survey duration on detection efficiency was observed. As suggested in previous research, detection increased when the survey lasted longer (Pierce and Gutzwiller 2004; Gooch et al. 2006). This reinforces the importance of standardised surveys, where the recording of survey time is essential to interpret the results. Contrary to data reported by Pierce and Gutzwiller (2004), survey duration had no significant effect on the accuracy of distance estimations but, again, the distance estimation task was quite simple.

There was a facilitative effect of group size on performance. The number of models detected initially increased with group size but then leveled off, with no statistically significant difference between groups of four and eight participants. As groups grew larger,
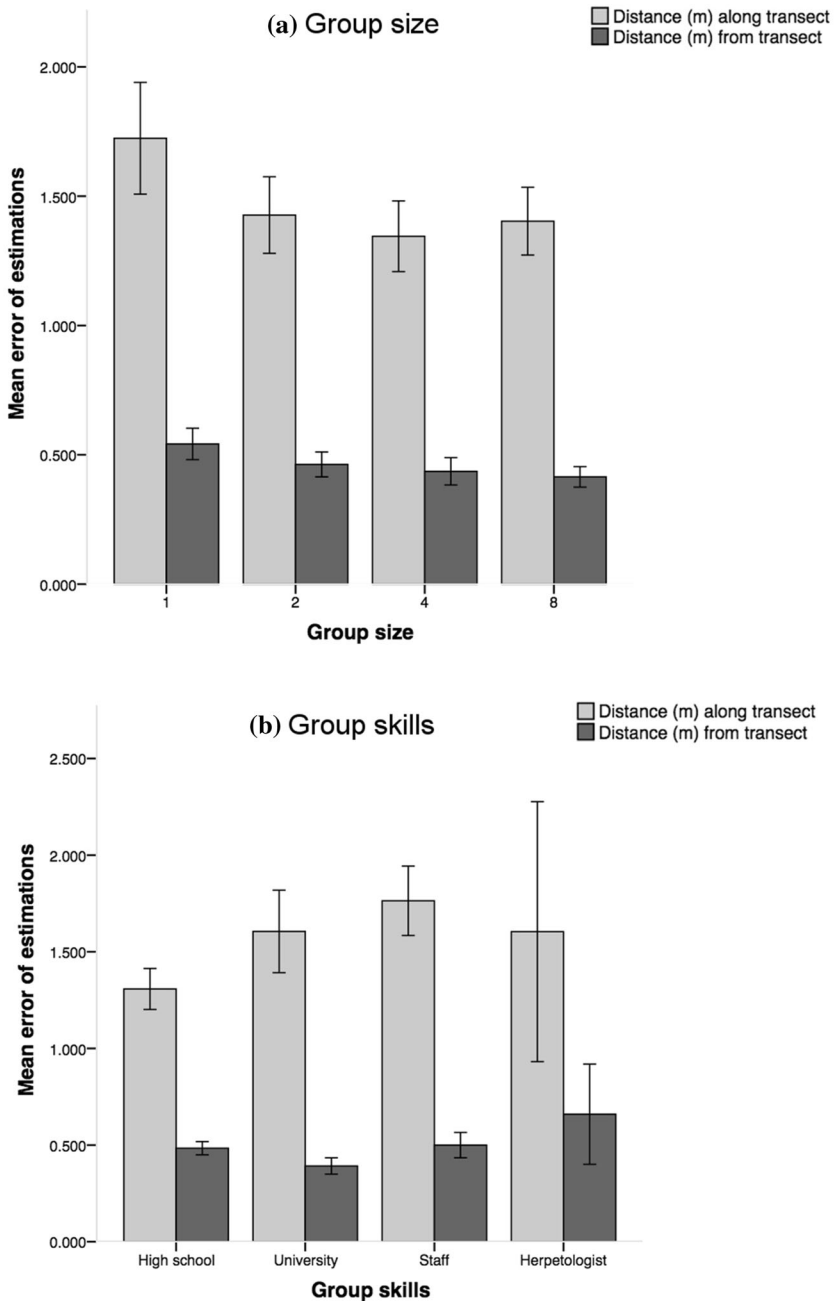
Fig. 6 Mean error in estimations of perpendicular distance from and distance along the transect for **a** different group sizes, **b** different group skills when analyzing the group dataset. *Error bars* are 1 *SE*

participants may have been able to share task responsibility and focus on different sections of the footpath. Personal observations during the trials suggested that distraction might be the cause of the similarity in detection ability of the larger groups, as with an increased number of people, disturbance and interference between members increased. Distance estimation accuracy, however, did not seem to be associated with a greater number of observers.

The experiment disclosed significant differences in the positions of models detected and also in the proportion of 'species' detected. The experiment was designed to replicate real field conditions. Amphibians and reptiles in Cusuco National Park are terrestrial as well as arboreal, the reason for which models were placed at different heights in the canopy. Despite being equally distributed in the three height categories, models were found more often in the eye-to-knee level. Different factors might have affected the detection rate at different heights, including foliage density, light availability and participant expectations to find models clinging to the vegetation.

One might reasonably ask how the data conform to models of detection performance that are incorporated into population density software such as Distance (Thomas et al. 2010). Although the purpose of this study was not to estimate species density, it is instructive to examine the data in an exploratory fashion to determine if there might be implications for such endeavours. The probability of detection for two skill levels (high school students and scientists of group size 1) is shown as a function of distance from the transect in Fig. 7.
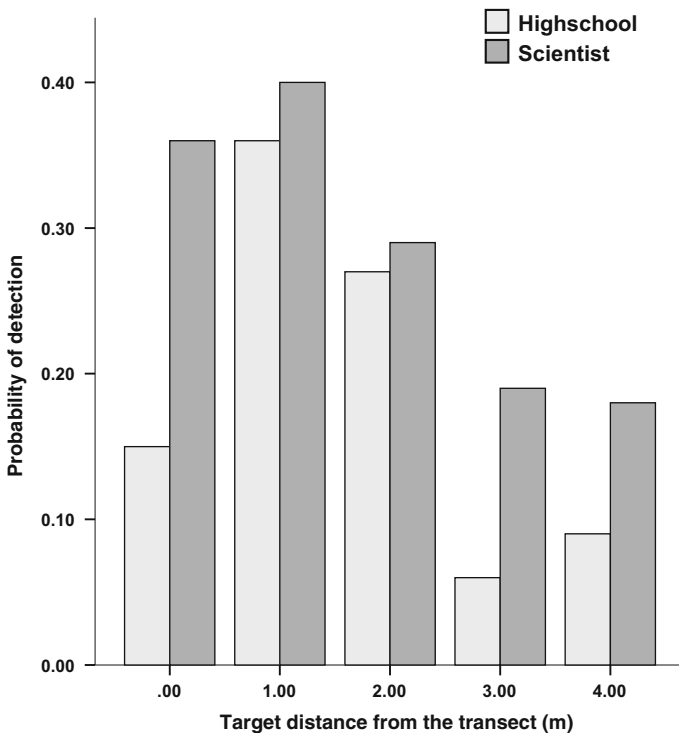

Fig. 7 Probability of detection as a function of distance for two skill levels (group size = 1)

A cursory examination of these data indicate that the assumption of perfect detection on the transect is violated by both groups. This can be accommodated in programs such as Distance (Cassey and McArdle 1999). The more challenging observation is that the function that can be best fitted to the data is very different for these groups. For example, a second-order polynomial function fits the data of scientists with an r^2 of .85. The same function produces an r^2 of only .53 for the high school students. At a less statistical level, it is clear that the decline in detection with distance is much greater for high school students than for scientists. Any estimates of true density must then take into account both skill level and group size in order to arrive at unbiased conclusions regarding species numbers in a region.

Limitations

From an analytic point of view, a limitation of this study was that we were unable to obtain data from a large number of participants in the more experienced groups; less experienced people were more numerous than experts. This is a common occurrence in all domains, including field ecology, and a consequence of the fact that in such projects there is always a large disparity in numbers of experts and novices. Additionally, local guides conducted the experiment with a slight difference in the protocol, which might have affected the results on distance estimations. It would be optimal to have data collected using the same procedures throughout all groups.

The models used did not represent species commonly seen in the field, which could have disadvantaged those experts who have very specific strategies for very specific target species with which they are familiar. It would therefore be useful to replicate this work using models that more closely resemble the field target species.

Conclusions

This study supports the view that when working in groups, non-specialist volunteer researchers can perform simple tasks and collect data as proficiently as more experienced observers with regard to object detection in complex natural habitats. In our study an optimum number of volunteers hovered around four individuals. This is not to suggest, however, that all observers are identical. In fact, some experienced observers were remarkable in their ability to detect models and make accurate distance estimations. We also acknowledge that experts are essential for identifying species, as well as training and leading the volunteers.

As the experiment focused mainly on detection ability, in confirming that volunteers can be as capable as their experienced counterparts in collecting reliable data for a baseline herpetology visual survey, this study reinforces the view that novice volunteers are able to bring valuable contributions to field research, not only financially, but also in practical terms. In so doing, it strengthens the idea that voluntourism expeditions can play an important role in global conservation and research programmes (Pattengill-Semmens and Semmens 2003) by accelerating data collection.

Also highlighted is the potential value of involving local communities when conducting field studies. Local guides demonstrated an excellent ability to detect models, showing their value to contribute to field research, thereby reinforcing the notion that local expert knowledge is becoming increasingly important for field conservation projects (Starr et al.

2011) and validating the belief that local experts can be used for quantitative wildlife studies (Gilchrist et al. 2005). Moreover, the involvement of local community members can be extremely beneficial for the socio-economical sustainability of the projects (Andrianandrasana et al. 2005).

This study has shown how the characteristics of monitoring protocols can have important implications for detection probability and sampling efficiency. Survey duration and number of surveyors had a substantial impact on detection probability during the experiment. Group size in particular appeared to be positively correlated to the increase in detection. This relationship suggests that examining and eventually adjusting these elements of survey protocols would improve the sampling efficiency of the research. For instance, in our results an optimum number of volunteers "per survey group" hovered around four individuals.

This information is extremely important for the development of long-term monitoring programs (Crouch III and Paton 2002) and for the design of studies involving volunteers (Foster-Smith and Evans 2003).

Drawing on these findings, the following recommendations can be made for managers planning future research conservation voluntourism works. The performance of volunteers in collecting data for monitoring studies should be evaluated. This should be done not only to compare volunteers to their professional counterparts and to evaluate overall data validity, but also to improve the protocols used for data collection. A thoughtful analysis and management of resources available can enable sampling efforts to be optimized and the efficiency of such studies to be improved. Volunteer training clearly contributes significantly to the success of monitoring programs (Genet and Sargent 2003), and it is strongly recommended that detection probabilities be incorporated into survey design and analysis in order to improve the accuracy of wildlife population estimation.

References

Alldredge MW, Simons TR, Pollock KH (2007) A field evaluation of distance measurement error in auditory avian point count surveys. J Wildl Manag 71(8):2759–2766

Andrianandrasana HT, Randriamahefasoa J, Durbin J, Lewis RE, Ratsimbazafy JH (2005) Participatory ecological monitoring of the Alaotra wetlands in Madagascar. Biodivers Conserv 14(11):2757–2774

Brightsmith DJ (2008) Rainforest expeditions and Earthwatch as funding partners for Macaw (Ara spp.) research in southeastern Peru. Neotrop Ornithol 19:173–181

Brown S, Morrison A (2003) Expanding volunteer vacation participation-an exploratory study on the mini-mission concept. Tour Recreat Res 28(3):73–82

Cassey P, McArdle BH (1999) An assessment of distance sampling techniques for estimating animal abundance. Environmetrics 10(3):261–278

Cohn JP (2008) Citizen science: can volunteers do real research? Bioscience 58(3):192–197

Crouch WB III, Paton PW (2002) Assessing the use of call surveys to monitor breeding anurans in Rhode Island. J Herpetol 36(2):185–192

de Solla SR, Shirose LJ, Fernie KJ, Barrett GC, Brousseau CS, Bishop CA (2005) Effect of sampling effort and species detectability on volunteer based anuran monitoring programs. Biol Conserv 121(4):585–594

Doan TM (2003) Which methods are most effective for surveying rain forest herpetofauna? J Herpetol 37(1):72–81

Dodd CK, Dorasio RM (2004) Using counts to simultaneously estimate abundance and detection probabilities in a salamander community. Herpetologica 60(4):468–478

Fewster RM, Buckland ST, Burnham KP, Borchers DL, Jupp PE, Laake JL, Thomas L (2009) Estimating the encounter rate variance in distance sampling. Biometrics 65(1):225–236

Fitzpatrick MC, Preisser EL, Ellison AM, Elkinton JS (2009) Observer bias and the detection of low-density populations. Ecol Appl 19(7):1673–1679

Foster-Smith J, Evans SM (2003) The value of marine ecological data collected by volunteers. Biol Conserv 113(2):199–213

Freilich JE, LaRue EL Jr (1998) Importance of observer experience in finding desert tortoises. J Wildl Manag 62:590–596

Genet KS, Sargent LG (2003) Evaluation of methods and data quality from a volunteer-based amphibian call survey. Wildl Soc Bull 31:703–714

Gilchrist G, Mallory M, Merkel F (2005) Can local ecological knowledge contribute to wildlife management? Case studies of migratory birds. Ecol Soc 10(1):20

Gooch MM, Heupel AM, Price SJ, Dorcas ME (2006) The effects of survey protocol on detection probabilities and site occupancy estimates of summer breeding anurans. Appl Herpetol 3:129–142

Hochachka WM, Martin K, Doyle F, Krebs CJ (2000) Monitoring vertebrate populations using observational data. Can J Zool 78(4):521–529

Holt BG, Rioja-Nieto R, Aaron MacNeil M, Lupton J, Rahbek C (2013) Comparing diversity data collected using a protocol designed for volunteers with results from a professional alternative. Methods Ecol Evol 4(4):383–392

IBM Corp. Released (2013). IBMS SPSS statistics for Mac, Version 22.0. IBM Corp, Armonk

Le Saout S, Hoffmann M, Shi Y, Hughes A, Bernard C, Brooks TM, Bertzky B, Butchart SHM, Stuart SN, Badman T (2013) Protected areas and effective biodiversity conservation. Science 342(6160):803–805

Léopold M, Cakacaka A, Meo S, Sikolia J, Lecchini D (2009) Evaluation of the effectiveness of three underwater reef fish monitoring methods in Fiji. Biodivers Conserv 18(13):3367–3382

Lind AJ, Welsh HH Jr, Tallmon DA (2005) Garter snake population dynamics from a 16-year study: considerations for ecological monitoring. Ecol Appl 15(1):294–303

MacKenzie DI, Kendall WL (2002) How should detection probability be incorporated into estimates of relative abundance? Ecology 83(9):2387–2393

Maxwell SE, Delaney HD (2003) Designing experiments and analyzing data: a model comparison perspective, vol 2. Taylor & Francis, New York, pp 319–343

McCarthy MA, Moore JL, Morris WK, Parris KM, Garrard GE, Vesk PA, Rumpff L, Giljohann KM, Camac JS, Bau SS (2013) The influence of abundance on detectability. Oikos 122(5):717–726

Newman C, Buesching CD, Macdonald DW (2003) Validating mammal monitoring methods and assessing the performance of volunteers in wildlife conservation "Sed quis custodiet ipsos custodies?". Biol Conserv 113(2):189–197

Olivet CR, Asquith N (2004) Ecosystem profile: Northern region of the Mesoamerica Biodiversity hotspot. Belize, Guatemala, Mexico Conservation International, Mexico and Central American Program

Pattengill-Semmens CV, Semmens BX (2003) Conservation and management applications of the reef volunteer fish monitoring program. Environ Monit Assess 81:43–50

Pierce BA, Gutzwiller KJ (2004) Auditory sampling of frogs: detection efficiency in relation to survey duration. J Herpetol 38(4):495–500

Pipino LL, Lee YW, Wang RY (2002) Data quality assessment. Commun ACM 45(4):211–218

Ryan TJ, Philippi T, Leiden YA, Dorcas ME, Wigley TB, Gibbons JW (2002) Monitoring herpetofauna in a managed forest landscape: effects of habitat types and census techniques. For Ecol Manag 167(1):83–90

Schmeller DS, Henry P, Julliard R, Gruber B, Clobert J, Dziock F, Lengyel S, Nowicki P, Deri E, Budrys E (2009) Advantages of volunteer-based biodiversity monitoring in Europe. Conserv Biol 23(2):307–316

Schmitt E, Sullivan K (1996) Analysis of a volunteer method for collecting fish presence and abundance data in the Florida Keys. Bull Mar Sci 59(2):404–416

Shirose LJ, Bishop CA, Green DM, MacDonald CJ, Brooks RJ, Helferty NJ (1997) Validation tests of an amphibian call count survey technique in Ontario, Canada. Herpetologica 53:312–320

Silvertown J (2009) A new dawn for citizen science. Trends Ecol Evol 24(9):467–471

Starr C, Nekaris KAI, Streicher U, Leung LKP (2011) Field surveys of the Vulnerable pygmy slow loris Nycticebus pygmaeus using local knowledge in Mondulkiri Province, Cambodia. Oryx 45(01):135–142

Stauffer HB, Ralph CJ, Miller SL (2002) Incorporating detection uncertainty into presence-absence surveys for marbled murrelet. In: Samson FB (ed) Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, DC

Thomas L, Buckland ST, Rexstad EA, Laake JL, Strindberg S, Hedley SL, Bishop JRB, Marques TA, Burnham KP (2010) Distance software: design and analysis of distance sampling surveys for estimating population size. J Appl Ecol 47(1):5–14

Wiest W, Bell B (1985) Stevens's exponent for psychophysical scaling of perceived, remembered, and inferred distance. Psychol Bull 8(3):457–470