Evaluating the Performance of Volunteers in Mapping Invasive Plants in Public Conservation Lands

Rebecca C. Jordan · Wesley R. Brooks ·
David V. Howe · Joan G. Ehrenfeld

Abstract Citizen science programs are touted as useful tools for engaging the public in science and for collecting important data for scientists and resource managers. To accomplish the latter, it must be shown that data collected by volunteers is sufficiently accurate and reliable. We engaged 119 volunteers over three years to map and estimate abundance of invasive plants in New York and New Jersey parklands. We tested their accuracy via collected pressed samples and by subsampling their transect points. We also compared the performances of volunteers and botanical experts. Our results support the notion that volunteer participation can enhance the data generated by scientists alone. We found that the quality of data collected might be affected by the environment in which the data are collected. We suggest that giving consideration to how people learn can not only help to achieve educational goals but can also help to produce more data to be used in scientific study.

Keywords Citizen science · Invasive plants · Training · Volunteers · Parklands · Monitoring program

Introduction

The ability of environmental managers to collect and analyze data across large temporal and spatial scales is clearly limited, especially given the current conditions of rapid environmental change. A trained cadre of motivated volunteers can be a means to further the environmental workforce. It is clear that there is potential for these projects to serve as a means of engagement for the public as well as an opportunity for scientists and resource managers to expand their capacity to collect data.

Citizen science programs are projects that enable public participation in research (Bonney and others 2009). Scientists often initiate these programs and they can result in data that are published in the primary, peer-reviewed literature. In some cases, citizens are considered technical participants and in others, these individuals have the opportunity to participate in decision-making. There are now numerous examples of citizen science programs, but some of the more prominent ones include Neighborhood Nest Watch (Evans and others 2005) and The Birdhouse Network (Brossard and others 2005) both initiated by the Cornell Lab of Ornithology.

While benefits to scientists in participating in citizen science have not been systematically explored, incentive for scientists and resource professionals to engage volunteers in citizen science can stem not only from the desire to grow a dataset beyond that in which they are capable of collecting themselves, but also in a desire to have a broader impact. Participants can develop their understanding of the underlying science of the issue (see review in Krasny and Bonney 2005; McCormick and others 2003; Jordan and others 2011). At a minimum, by simply providing the observational tools necessary for participation, citizen science involvement usually results in increased awareness of scientific processes (Pattengill-Semmens and Semmens 2003; Nerbonne and Nelson 2004) and can promote more advanced science process skills when volunteers are given opportunities to practice and develop them.

Broader social benefits are also possible. With respect to environmental issues, research has shown citizen science

R. C. Jordan (✉) · W. R. Brooks · D. V. Howe · J. G. Ehrenfeld
Department of Ecology, Evolution, & Natural Resources, School of Environmental and Biological Sciences, Rutgers University, New Brunswick, NJ 08901, USA
e-mail: jordan@aesop.rutgers.edu

initiatives can successfully promote civic engagement (Weber 2000). In a review of biological monitoring programs, Nerbonne and Nelson (2004) found several programs that resulted in increased civic awareness, involvement in local issues, and the creation of scientific data sets. Citizen groups have played a role even in shaping environmental policy (Dunlap 1992; Bierele and Cayford 2002). This enhanced efficacy might even translate into further environmental action-oriented behavior (e.g., Marcinkowski 1993).

Even given these potential benefits of volunteer participation in research, is it worthwhile for resource professionals and scientists to invest in the development of this type of programming? Are the data usable the data usable by resource professionals and scientists? Certainly concern over data quality exists and not all tasks are necessarily suitable for volunteers (e.g., Darwall and Dulvy 1996). Accuracy has been tested through the use of professional validation of either the sampling sites or pressed specimens in a number of environmentally oriented citizen science studies ranging across taxa and results have been variable. We describe some of these studies below.

In a study where volunteers were involved in identifying tree species, volunteers and professionals were able to generate comparable results at the generic level but when species-level data were analyzed volunteer accuracy was considerably less than that of professionals (Brandon and others 2003). In another study in which volunteers engaged in a similar task, volunteer and expert accuracy not only varied by identification but also by data collecting tasks (Bloniarz and Ryan 1996). In a study investigating volunteer accuracy in both detecting invasive plants and taking additional measures, volunteers were found to be accurate ($\sim$70%) when compared to experts ($\sim$85%; Crall and others 2011).

Accuracy may differ by taxa as well. While plant identification (above) yielded variable results, two invertebrate studies (molluscs and crabs respectively) yielded higher accuracy among volunteers when compared with experts (Thelen and Thiet 2008; Delaney and others 2008). In the former, volunteers were as accurate as professionals and in the latter the amount of education was shown to be an important variable when predicting volunteer accuracy and that undergraduate students were highly accurate (95% by one measure). In that study, data from seventh graders were found to be 80% accurate. In contrast, a study of hemlock woolly adelgid identification yielded volunteer accuracy level to be much lower than that of professionals (Fitzpatrick and others 2009). Variable results have also been reported for birds and anurans with some tasks yielding higher accuracy than others (e.g., Pierce and Gutzwiller 2007; Dickinson and others 2010 respectively).

While not all datasets are clearly useful for resource professionals and scientists, many researchers have argued

that when the task is specifically designed for volunteers and sampling efforts adjusted accordingly (e.g., Boudreau and Yan 2004, early invasive species detection), volunteers can collect highly accurate data. In addition, if volunteers are given the chance to improve over time, accuracy can increase (e.g., Schmeller and others 2009; reported volunteer improvement with time). Additionally, if certain skills are targeted such as the development of a search image (Humphreys 1989), obtaining accuracy may be more efficient. Such training would need to be considered prior to the training period. Demands for precision and accuracy should be taken into account as decisions about volunteer specialization and number are made (e.g., Firehock and West 1995). In addition, post-processing of the data needs to be seriously considered (Link and Sauer 1999; Dickinson and others 2010; Crall and others 2011).

In this paper, we examine the efficacy of using volunteers in an invasive plant mapping study conducted in public conservation lands in New York and New Jersey from 2006 to 2008. We describe how we were able to collect and validate data to be used both in peer-reviewed publication and to inform policy-makers. It is the combination of validation approaches that makes our study unique and allows us to draw conclusions about the different aspects of the data collection tasks. We use this context to highlight critical issues in the use of lay volunteers in environmental monitoring.

Project Context and Research Questions

Invasive plants pose a serious threat to native ecosystem biodiversity. Although the presence of many invasive species in the northeastern United States is known, knowledge of their distributions is scant (Luken 2003). This deficit limits the ability of natural resource managers and policy-makers to respond to issues concerning invasive plants and to establish policies and regulations intended to limit their further spread. The impetus for the project described here was to learn more about invasive plants in forested parkland in the Highlands physiographic province along the border between New Jersey and New York (Fig. 1). The area includes Harriman and Bear Mountain State Parks in New York, and a series of parks in the Ramapo Mountains of New Jersey, as well as other nearby protected lands. A major goal of the project was to test the hypothesis that trails were positively associated with invasive plant species distribution. We, therefore, needed to characterize plant distribution on-trail to compare to that off-trail. While off-trail transects were to be completed by trained graduate students, help was necessary to cover the on-trail trail transects. We determined that to map a sufficient subsample, we would require surveying along well
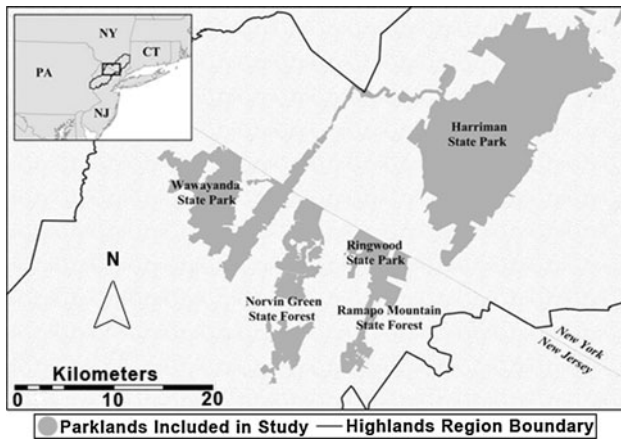
Fig. 1 Map indicating the Highlands Province and the approximate study area

over 150 miles of transects. Thus, the idea of using volunteers was appealing for logistic purposes, but also opened up the project to testing hypotheses about volunteer accuracy which are the focus of this paper: (a) volunteers can be trained to correctly identify invasive plants of interest, (b) volunteer data can be highly accurate, (c) variation among volunteer observations would be low, and (d) volunteers and professionals would differ in their accuracy.

Methods

Volunteer Recruitment and Training

During February 2006, 2007, and 2008, we recruited 58, 35, and 26 volunteers respectively from the New York-New Jersey Trail Conference (NYNJTC); a recreational hiking association with a membership of about 10,000 individuals and about 100 clubs. These volunteers were recruited via an email flyer sent to the entire membership. We offered no material incentives. If volunteers were able to undertake the hiking and could attend the training sessions, they were accepted. Volunteers wishing to return during the years following their participation were allowed to do so, but because of a likely increase in their ability, their data were excluded from the analysis.

Volunteers attended one all-day training session led by the authors in early June and a follow-up 'debriefing' session after they had collected their data in early July. During the training session, we gave participants background information about the ecology and impacts of invasive species. The participants then received hands-on species identification training in identifying a target list of invasive plants (22 species in 2006, but given the rarity of plants spotted, 12 in 2007, and 13 in 2008). We categorized

target plants as trees, shrubs, vines, and herbs and we instructed volunteers to scan in a stratified manner from canopy to ground. We provided the volunteers with a field ID guide specific to the project. We also trained participants in a field-based semi-quantitative data collection protocol.

Invasive Plant Collection

Volunteers collected data in pairs. We assigned each pair a 2 mile ($\sim$3.2 km) length of trail to survey. Volunteers surveyed approximately 150 miles ($\sim$241 km) of trails. The protocol involved pairs of volunteers hiking the assigned stretch of trail, stopping every 0.1 mile (0.16 km) to record presence and abundance of target species and to collect samples when they first encountered a target species.

Volunteers scanned for the target species in two zones on both the right and left sides of the trail at each stop: 'trailside' and 'distant'. Volunteers estimated the trailside zone in 2006, but delineated it on three sides with lengths of rope in 2007 and 2008. The trailside zone was a 6 ft (1.83 m) $\times$ 20 ft (6.10 m) rectangle directly off the trail. The distant zone included the area directly back from the rectangle as far as observable (Fig. 2). We instructed volunteers not to leave the trail. Volunteers estimated relative abundance as 'few', 'some', or 'many'. Both the zone and abundance assignments required judgments subject to variation among observers.

For sample collection, we provided volunteers with a series of bags and labeling stickers for use while on the trail, and a plant press for sample preservation. We trained them in the procedures to label and preserve the samples. Volunteers returned samples in the press.
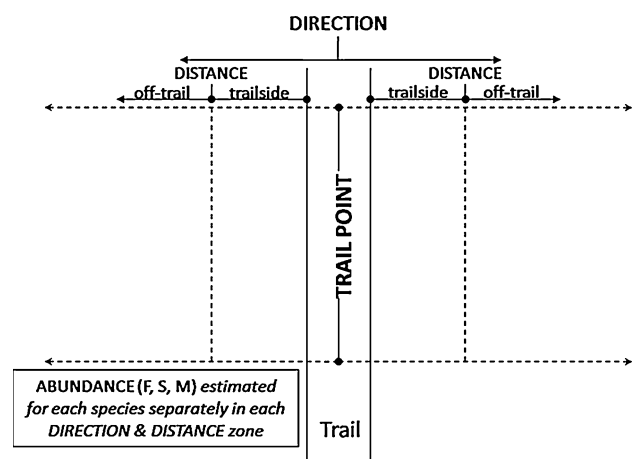
Fig. 2 Diagram of the data collection plots. Note the presence of a trailside zone (1.83 $\times$ 6.10 m)

Reliability of Volunteer Data

We used multiple methods to assess the accuracy of volunteer-generated data: (1) Pressed samples to test species recognition, (2) trail-point validation to measure volunteer accuracy on trails, (3) common trail data collection to measure the repeatability of volunteer-generated data, and (4) common trail data collection to compare the performances of volunteer and 'professional' recruits.

1) Pressed Samples

Project personnel assessed specimen accuracy and we report these results as percentages.

2) Trail Point Validation

To assess volunteer ability to detect target plants and estimate their abundance along trails, we compared data collected by volunteers at a subset of points along their trails with data collected by three pairs of specially trained staff, henceforth referred to as 'validators'. These personnel repeatedly assisted with training and spent more time on the trails than other staff. Thus, they have well-developed and specialized skills for identifying our target plants and in applying the protocol. We also found high repeatability in validator-generated data and this coupled with the much higher variability in volunteer data gives us confidence that we can use the validator-generated data as an acceptable standard by which to measure accuracy. 30% of volunteer-collected data points were validated. In 2007 and 2008, roughly 50% of the data points were validated.

The subset of validated points included the starting point and the next ten points (i.e., 50%) n of their trail, which volunteers had marked using flags and flagging tape. We always used the first ten points in which the volunteers were likely still new to the data collection task. It is possible that with each point volunteers could have improved. While post-hoc analysis indicated that there were no significant differences from the first versus the tenth point in volunteer accuracy, we wanted to make the most conservative accuracy estimate.

Validators went out to assess accuracy within two weeks of volunteer submission of data. We assessed volunteer accuracy at two levels. The first level considered the presence or absence of a plant at a trail location without regard to either of the other parameters, i.e. zone or abundance. For this analysis, an observation of a species was marked as a true positive (TP) if both volunteer and validator confirmed the plant, a true negative (TN) if both volunteer and validator agreed that the plant was not present, a false negative (FN) if the validator noted the plant but the volunteer did not, or a false positive (FP) if the volunteer noted a plant but the validator did not. We summed the observation types for each year and we calculated percentages of total observations for each category. We also calculated an accuracy score using this formula: $100 \times TP/(TP + FP + FN)$. We excluded TN observations because their high rate could mask variation. We averaged these scores within the zone and abundance variables to see if these parameters affected accuracy. We also conducted a Spearman's Rank Correlation test using these scores and the number of invasive species found on a given trail to look for an effect of level of invasion on accuracy.

At the next level, we inspected true positives for agreement at each point in the trail with respect to zone and abundance parameters. 'Correct observations' include those with agreement between volunteer and validator on both the zone and abundance parameters which are described below. 'Abundance errors' include observations where there was agreement between volunteers and validators regarding the zone of a species but disagreement regarding abundance estimates. 'Zone errors' include observations where volunteer and validator disagreed as to whether a species fell in the trailside zone or the distant zone. Disagreements where either volunteers or validators noted a species at both levels of zone, but the other noted the species in only one level, we assign to the 'Ambiguous error' category. These could be errors in estimation of zone, wherein the volunteers have incorrectly allotted some plants to either inside or outside of the trailside zone, or these errors could be false positives or negatives. We calculated percentages of each of these categories for each year. Note that the total percent for a year can exceed 100% because a true positive as we have defined it can contain more than one error type. With such limited data of this nature, it is hard to say how different percentages would need to be in order for meaningful conclusions to be drawn. We therefore report percentages without inferential statistics.

3) Common Validation Trail

We examined the repeatability of volunteer-produced data. To do this, we had all volunteers collect data along 1 mile (1.6 km) 'validation' trails that were explicitly marked at 11 points each. We used separate trails for NJ and NY volunteers in 2006 and 2008, but a single trail for convenience in 2007. The most experienced validator team also surveyed these trails and we used their data as the standard. We calculated an accuracy score as described above for each volunteer pair for their efforts on both the validation trail and their experimental trail.

For all five trails, we plotted the scores from the validation trails to look at variation among volunteers. We inspected these data separately because the validation trails varied from one to another. The variation in validation trails resulted from the desire to place the trails not only

near the data collection sites but also near the roadside for convenience. Qualitatively, the 2006 NJ trail was heavily invaded while the NY trail, though heavily vegetated, included considerably fewer invasive species. The 2007 and 2008 trails, however, were intermediate in level of invasion and general vegetation. In addition, we looked for correlation between volunteer accuracy scores using Spearman's Rank Correlation. Given the extreme nature of the 2006 validation trails, we excluded these data from our validation analysis.

4) Professionals

Last, we wanted to see how the performance of specially trained volunteers might compare with the performance of professionals who had strong botanical backgrounds. We did not, however, give these professionals explicit training on the target species. To accomplish this, in 2006 eight botanically oriented professionals, who were not project personnel, individually surveyed the New Jersey validation trail. We provided the professionals with the same materials as the volunteers but not the protocol training. Again, the most experienced validation team's data were used to determine accuracy. We generated percentages of observation types and error categories as described above for both the volunteers and professionals. We also calculated mean accuracy scores as described above for the two groups. Note that the purpose of calculating these scores was not only to determine the extent to which target species were identified but also to determine variation in zone and abundance, which is subject to uncertainty in both judgment and adherence to the protocol. Given the low sample size and variability in the two groups (i.e., 17 volunteer teams and eight professional teams), we chose not to conduct inferential statistics.

Results

In general, volunteers were highly accurate in identifying their pressed samples. Volunteers were highly accurate overall (on average 97.2% for all three years), yet when sites without invasive plants were removed from the dataset, accuracy was considerably lower. In addition, when the data were partitioned by zone and unknown factors, accuracy was also lower. Volunteer accuracy could not be predicted by performance on the validation trail. Finally, volunteer accuracy was about 15% lower than that of professionals.

Pressed Samples

Volunteers largely identified the samples they collected correctly (Table 1). We found some variation in accuracy among species. Volunteers in all years always correctly identified three of the 13 target species, but only about half of the samples reported as Norway maple, *Acer platanoides*, and mile-a-minute weed, *Persicaria perfoliata*, were correct. We found little variation in identification accuracy of each species across years with the exception of *Acer platanoides*, and *Persicaria perfoliata*.

Trail Point Validation

Volunteer TP + TN rates were 98.4, 98.6, and 94.6% for 2006, 2007, and 2008, respectively (Table 2). Volunteer

Table 1 Percent accuracy of pressed sample identification and the number of samples pressed (*n*) by year; also shown is the number of trail points at which validators found the species

Species	No. of trail points	Overall accuracy	2006		2007		2008	
			n	Accuracy	*n*	Accuracy	*n*	Accuracy
Berberis thunbergii	343	100	27	100	18	100	17	100
Microstegium vimineum	218	86.5	22	86.4	–	n/a	15	86.7
Rosa multiflora	85	95.5	18	100	15	100	11	81.8
Alliaria petiolata	81	100	22	100	13	100	10	100
Celastrus orbiculatus	79	97.3	13	92.3	15	100	9	100
Rubus phoenicolasius	48	93.9	10	90.0	15	100	8	87.5
Ailanthus altissima	47	95.0	13	92.3	4	100	3	100
Aralia elata	37	90.5	8	87.5	2	100	11	90.9
Euonymus alata	30	96.6	11	100	17	100	1	0.0
Acer platanoides	24	59.1	13	84.6	3	0	6	33.3
Persicaria perfoliata	24	42.9	2	50.0	3	0	2	100
Cynanchum louiseae	14	100	–	–	1	100	5	100
Lonicera japonica	6	87.5	8	100	3	66.7	5	80.0

accuracy increased with plant abundance and when plants were in the trailside rather than the distant zone (Table 3). We also found that accuracy was positively correlated with the number of invasive species found on the trail (Spearman's Rank Correlation, $\rho = 0.548$, $n = 51$, $P < 0.001$). In examining the TP observations, we found that 45.5, 50.0, and 46.0% in 2006, 2007, and 2008, respectively, were correct in both the zone and abundance parameters (Table 4). We found abundance errors in 25.0, 37.9, and 32.5% of TP observations over the three years, and we found zone errors in 7.6, 10.4, and 5.3% of observations. The ambiguous error rates were 29.5, 12.1, and 21.5%.

Common Validation Trail

We found no correlation between volunteer accuracy on experimental and validation trails (Spearman's $\rho = 0.248$, $n = 41$, $P = 0.118$) and accuracy scores on the validation trails varied considerably among volunteers (Fig. 3). The mean accuracy score was 53.7 for all volunteers.

Professionals

Professionals had higher numbers of spotting invasive plants, with a TP rate of 17.8 compared to 14.2 for volunteers, however, professionals made slightly more FP errors. Of the TP observations, 44.6% were without either an abundance or zone error for the professionals, compared to 35.8% for the volunteers. The accuracy of the professionals generally exceeded the accuracy of the volunteers (Table 5). The mean accuracy scores were 69.4 for the professionals and, as stated above, 53.7 for the volunteers.

Table 2 Percentage of observations from volunteer presence-absence data of experimental trails in each analysis category by year

	2006	2007	2008
True positives	2.22	1.19	5.59
True negatives	96.16	97.42	88.97
False negatives	1.09	0.74	2.63
False positives	0.53	0.65	2.82

**Table 3** Accuracy scores averaged within the zone and abundance variables

	Accuracy scores		
	Few	Some	Many
Trailside	50.0	75.2	87.8
Distant	31.3	48.7	68.5

Discussion

In this paper, we tested several hypotheses regarding whether volunteers could gather accurate ecological data. We applied a variety of approaches to answer these questions, and we found in general that, volunteers can correctly identify a number of invasive plant species and that they could collect accurate data. Variability among volunteers was not high on our experimental trails and we

Table 4 Percentages of fully correct observations collected by volunteers and error types by year

	2006	2007	2008
Correct observations	45.5	50.0	46.0
Abundance errors	25.0	37.9	32.5
Zone errors	7.6	10.4	5.3
Ambiguous errors	29.5	12.1	21.5
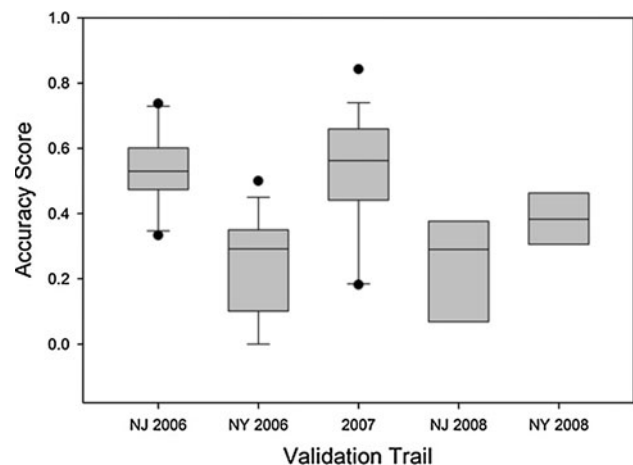
Fig. 3 Boxplot of accuracy scores on the validation trails. $N = 6, 17, 16, 4$, and 8 teams of volunteers, respectively

Table 5 Accuracy scores of volunteers and professionals on the 2006 New Jersey validation trail

	Volunteers	Professionals
True positives	14.2 ± 3.53	17.8 ± 1.47
Correct observations	35.8 ± 14.11	44.6 ± 9.88
Abundance errors	35.8 ± 16.08	30.2 ± 10.07
Zone errors	8.2 ± 5.18	4.4 ± 4.08
Ambiguous errors	28.4 ± 8.94	25.2 ± 9.08
True negatives	73.7 ± 3.42	74.3 ± 1.77
False negatives	9.3 ± 2.72	5.1 ± 1.25
False positives	2.6 ± 2.93	2.8 ± 1.25
Accuracy scores	53.7	69.4

Standard deviations are provided. There were 17 volunteer teams and eight professional teams

found volunteer performance to be slightly less than that of untrained-on-our-protocol professionals. Most importantly, however, <mark>our data support the notion that volunteer participation can enhance the data generated by scientists alone.</mark> The combined trail length surveyed was clearly beyond that which project personnel alone could have completed and the accuracy of volunteers approached that of professionals. Several issues regarding the extent to which volunteers could collect accurate data in varying contexts, however, warrant consideration.

Pressed Samples

We instructed volunteers to save and press a sample on their first encounter of a target species so we could test their ability to identify plants. The materials we provided to collect samples may have been overly burdensome to some volunteers, but perhaps more importantly, volunteers were more concerned with correctly applying the data collection protocol and accurately identifying the plants. Some volunteers likely viewed the plant pressing procedures as adjunct to the main task, and we might expect this, given how we allotted time and effort in the training.

The pressed samples we received, however, indicate that the volunteers were highly accurate in identifying plants, although some plants might be more difficult to learn than others. In addition, accuracy did not correlate with species abundance or the likelihood that a volunteer would encounter a species, meaning that certain plants might be more difficult to identify than others in spite of additional time spent encountering these plants. For example, Japanese barberry, *Berberis thunbergii*, was the most common and one of the most accurately identified species, while Japanese stilt grass, *Microstegium vimineum*, was also extremely abundant but often overlooked or misidentified.

Trail Point Validation

Along the experimental trails, volunteers were highly accurate, with the most common type of correct response being a true negative; meaning that both validator and volunteer agreed that none of the experimental plants were present. Along these trails, it is unclear the extent to which other plants (i.e., not those targeted by this study), whether native, exotic, or invasive, were present, but in general, volunteers were instructed not to try to identify every plant at the site, but rather to find those that were of interest to our study. In the presence of many other plants, therefore, a major challenge might be to discern study plants from background noise.

Given that true negatives had an inflationary effect on volunteer accuracy rates, we can then ask: to what extent do volunteers correctly identify plants that are present? In

analyzing the data from points where study plants were present, volunteers accurately identified plants ranging from 45 to 50% depending on the year. When making an error, volunteers on average were only slightly more likely to overlook a plant than to misidentify a particular plant as being one of interest to our study. Volunteers were more likely to find plants close to the trail and if the species was abundant. Furthermore, volunteer accuracy increased with the number of invasive plants found on the trail. This might reflect increased vigilance as more plants are spotted or simply experience. Many trail sites had no invasive plants and negative data can become monotonous. In support of this notion of increased vigilance with task difficulty is our finding of little difference in accuracy rates among years. In 2006, volunteers had twice as many plants to identify and yet they were less likely to identify a particular plant erroneously.

Once the volunteers had correctly identified a plant of interest, they then had to correctly assess zone and estimate abundance. Abundance errors were more common than zone errors, even considering that some errors of uncertain cause could indeed have been zone errors. For example, in some situations, both volunteer and validator found more than one plant of a particular species at a site, but the volunteer reported observing the species both inside and beyond the trailside zone. This could be a zone error; however, it could be a mistaken identification. An inverse situation where the validator found the species in both zones and the volunteer in only one could be attributable to a volunteer not recognizing that some plants crossed the boundary between zones, or it could have resulted from unseen plants.

Volunteers and the validator estimated abundance on a rough and rather subjective scale, from few to some to many. Using such a measure, disagreements are unavoidable. Further, there could be a temporal factor. While local abundance was unlikely to change during the week or two following the volunteer data collection period, it is possible that certain species could have become more visible during the growing period because of their own growth or because of changes in surrounding vegetation.

Common Validation Trail and Professional Accuracy

We feel our comparison of the performances of our volunteers and our 'professional' group indicates that the use of volunteers for collecting data is practicable. The variation in accuracy among volunteers was higher, but the difference in accuracy rates was not dramatic: 53.7% on average for volunteers compared to 69.4% for the professionals. Trends in error rates were similar. While it is hard to know to what extent these differences in error rate are significant, they are comparable to a very similar study

with volunteers from the Midwestern and Western United States (Crall and others 2011). In that study, accuracy varied by plant species with those being more difficult for professionals causing somewhat greater difficulty for volunteers. Yet, about half of the 12 targeted species resulted in volunteer accuracy rates over 75%. These were the most conspicuous species. In addition, volunteers correctly identified species 70% of the time when compared to professionals at 85%.

Given the roughly 30% error rate in professionals in our study, these results suggest that participants may need specific training for the project at hand and that one cannot assume that advanced knowledge in a particular area will convey greater proficiency than that of well-trained volunteers for a particular task. Therefore, if one were to employ professionals, resources would likely need to be allocated to the training of professionals.

The successes and failures of the volunteers engaged in this study provide insight for future training efforts for monitoring programs. With more specific and detailed data collection protocols, comes greater opportunity for data uncertainty. In some cases, rough but more reliable estimates might make sense. While our project had very specific goals for testing ideas about invasive species spread, other projects might seek to monitor the spread of invasive species. A different protocol might be more appropriate for early detection/rapid response type programs.

Nonetheless, we were able to collect data from which conclusions can be reliably drawn. For example, we know that in spite of the fact that the study sites were uniform in ecology (i.e., geology and climate) and in close proximity, the frequency of common invasive species (i.e., *Berberis thunbergii*, *Microstegium vimineum*, *Rosa multiflora*, *Celastrus orbiculatus*) was variable between study sites and within location (i.e., trailside or off-trail) (Ehrenfeld and others, unpublished data). Further we found that *B. thunbergii* was correlated with moderate density and off trail and with this, all species studied were found away from trail side (Ehrenfeld and others, unpublished data). Here, we conclude that multiple factors, including the presence of trails but also including land-use history, topography, and propagule pressure, are necessary to explain the distribution of these exotics in deciduous forested lands.

Implications

Important considerations might be the tensions that arise when considering the time allocated to aspects of the training program. Given limited time and resources, necessary trade-offs exist when considering both data quality and educational goals established for the volunteers either by scientists, managers, or agencies involved in the study.

First, the amount of detail incorporated into the training protocol can greatly affect the quality of data collected, but trainers must balance this with the available time to train any given volunteer. We designed our protocol to be quick, keep the volunteers on the trails for logistical purposes, and to collect a sufficient data set that would enable us to address our scientific questions. In an attempt to keep the method simple yet semi-quantitative, we introduced sources of uncertainty, e.g., the back line of the trailside zone had to be estimated, we tried the protocol without the ropes one year, and the difference between 'few', 'some' and 'many' is subjective. Further, this trade-off between simplicity of method, and accuracy and precision of data was unavoidable given our intention to gather data across a broad spatial scale that necessitated a larger number of volunteers. Trainers need to consider how much can be taught in a limited time, how willing volunteers are to spend time in training sessions, and the kind(s) of data needed. For example, if we required volunteers to attend two full days of training, we may not have had enough volunteers to run the project, even though the data might have been more accurate. Additionally, if we focused solely on plant identification, we would not have had enough information to test our ideas adequately. We needed to find a balance between the amount of information necessary and the accuracy and precision of those data.

A similar trade-off in time allocation could be considered for the education goals which may exist for the project. Before engaging in a complex training program one could ask, to what extent should time be spent teaching volunteers to transfer skills to varying contexts? Teaching for transfer necessitates preparing an individual to complete a task, recognize when the task is appropriate, and be able to change to accommodate a new context. This level of training often requires time for reflection, practice, and learning how to adapt a conceptual model of the task at hand (see Bransford and others 1999, for a more detailed explanation). Clearly, training for transfer will take more time than training to identify plants in a single context. Our project provided context-specific training in a very short time.

Moving beyond educational goals alone, the time allocated toward broader learning could have implications for data accuracy and precision as well. For example, given the context for our project, volunteers generated an accurate data set. We can attribute much of this accuracy, however, to large amounts of "no data" points. When thinking about the points with data, volunteers missed critical plants and often did not accurately estimate abundance. We do not attribute these errors to lack of learning the plants in a classroom setting, but rather from individuals overlooking plants in the backdrop of other vegetation. If more of the sites were invaded, would volunteers have generated a less accurate data set, or would the opportunity to engage with more target

plants have resulted in an increased ability to pick out plants from the background? Perhaps greater training for transfer would have enabled the volunteers to identify the plants in a greater variety of backgrounds (i.e., trail sites).

We recommend that project designers consider both the data quality and educational needs of the project when considering time allocations. While we were able to generate a reliable data set to address our questions, our volunteers likely would not have collected an accurate dataset in heavily invaded contexts or where abundance estimation is essential. Teaching these skills would involve further training and must allow more time or another means of tolerating error in the data set.

In conclusion, we found that volunteers can correctly identify invasive plant species and that they could collect accurate data. Variability among volunteers was minimal on the experimental trails and volunteer accuracy was slightly less than that of our professional audience. These data support our conclusions that volunteers can help generate data beyond that of which scientists can collect alone. Finally, we contend that inclusion of volunteers can enable broader societal benefits which may also be of importance for scientists and managers.

References

Bierele TC, Cayford J (2002) Democracy in practice: public participation in environmental decisions. RFF Press, Washington, DC, USA, p 148

Bloniarz DV, Ryan HDP III (1996) The use of volunteer initiatives in conducting urban forest resource inventories. Journal of Arboriculogy 22:75–82

Bonney R, Ballard H, Jordan R, McCallie E, Phillips T, Shirk J, Wilderman C (2009) Public participation in scientific research: defining the field and assessing its potential for informal science education. In: A CAISE inquiry group report center for advancement of informal science education. CAISE, Washington, DC. http://caise.insci.org/uploads/docs/PPSR%20report%20FINAL.pdf. Accessed 15 March 2010

Boudreau SA, Yan ND (2004) Auditing the accuracy of a volunteer-based surveillance program for an aquatic invader Bythotrephes. Journal of Environmental Monitoring 91:17–26

Brandon A, Spyreas G, Molano-Flores B, Carroll C, Ellis J (2003) Can volunteers provide reliable data for forest vegetation surveys? Natural Areas Journal 23:254–261

Bransford JD, Brown A, Cocking R (eds) (1999) How people learn: brain, mind, experience, and school. National Academy Press, Washington, DC, USA, p 374

Brossard D, Lewenstein B, Bonney R (2005) Scientific knowledge and attitude change: the impact of a citizen science project. International Journal of Science Education 27:1099–1121

Crall AW, Newman GL, Stohlgren TJ, Holfelder KA, Graham J, Waller DM (2011) Assessing citizen science data quality: an invasive species case study. Conservation Letters. doi:10.1111/j.1755-263X.2011.00196.x

Darwall W, Dulvy N (1996) An evaluation of the suitability of nonspecialist volunteer researchers for coral reef fish surveys, Mafia Island, Tanzania—a case study. Biological Conservation 78:223–231

Delaney DG, Sperling CD, Adams CS, Leung B (2008) Marine invasive species: validation of citizen science and implications for national monitoring networks. Biological Invasions 10:117–128

Dickinson JL, Zuckerberg B, Bonter DN (2010) Citizen science as an ecological research tool: challenges and benefits. Annual Reviews of Ecology, Evolution, and Systematics 41:149–172

Dunlap R (1992) Trends in public opinion toward environmental issues: 1965–1990. In: Dunlap R, Mertig A (eds) American environmentalism: the US environmental movement, 1970–1990. Taylor and Francis, Philadelphia, PA, USA, pp 89–116

Evans C, Abrams E, Reitsma R, Roux K, Salmonsen L, Marra P (2005) The neighborhood nestwatch program: participant outcomes of a citizen-science ecological research project. Conservation Biology 19:589–594

Firehock K, West J (1995) A brief history of volunteer biological water monitoring using macroinvertebrates. Journal of North American Benthological Society 14:197–202

Fitzpatrick M, Preisser E, Ellison A, Elkinton J (2009) Observer bias and the detection of low-density populations. Applied Ecology 19:1673–1679

Humphreys G (1989) Visual search and stimulus similarity. Psychology Review 96:433–458

Jordan RC, Gray ST, Howe DV, Brooks WR, Ehrenfeld JG (2011) Conservation education and citizen science: using an invasive plant case study to highlight issues of practice. Conservation Biology. doi:10.1111/j.1523-1739.2011

Krasny M, Bonney R (2005) Environmental education through citizen science and participatory action research. In: Johnson EA, Mappin MJ (eds) Environmental education or advocacy: perspectives of ecology and education in environmental education. Cambridge University Press, New York, USA, pp 292–319

Link WA, Sauer JR (1999) Controlling for varying effort in count surveys—an analysis of Christmas bird count data. Journal of Agricultural, Biological, and Environmental Statistics 4:116–125

Luken JO (2003) Invasions of forests in the Eastern United States. In: Gilliam FS, Roberts MR (eds) The herbaceous layer in forests of Eastern North America. Oxford University Press, New York, USA, pp 283–301

Marcinkowski T (1993) Assessment in environmental education. In: Wilke RJ (ed) Environmental education teacher resource handbook. Corwin Press, Thousand Oaks, CA, USA, pp 143–197

McCormick S, Brown P, Zavestoski S (2003) The personal is scientific, the scientific is political: the public paradigm of the environmental breast cancer movement. Sociological Forum 18:545–576

Nerbonne JF, Nelson KC (2004) Volunteer macroinvertebrate monitoring in the United States: resource mobilization and comparative state structures. Society and Natural Resources 17:817–839

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

Pierce BA, Gutzwiller KJ (2007) Interobserver variation in frog call surveys. Journal of Herpetology 41:424–429

Schmeller DS, Henry PY, Julliard R, Gruber B, Clobert J and others (2009) Advantages of volunteer-based biodiversity monitoring in Europe. Conservation Biology 23:307–316

Thelen BA, Thiet RK (2008) Cultivating connection: incorporating meaningful citizen science into Cape Cod national seashore's estuarine research and monitoring programs. Park Science 25(1):74–80. http://tinyurl.com/parkscience

Weber EP (2000) A new vanguard for the environment: grass–roots ecosystem management as a new environmental movement. Society and Natural Resources 13:237–259