Evaluating the Quality of Citizen-Scientist Data on Pollinator Communities

C. KREMEN,* K. S. ULLMANN,†,‡ AND R. W. THORP‡

*Department of Environmental Sciences, Policy and Management, 130 Mulford Hall, University of California, Berkeley, CA 94720-3114, U.S.A., email ckremen@berkeley.edu
†Xerces Society for Invertebrate Conservation, 4828 NE Hawthorne Avenue, Portland, OR, U.S.A.
‡Department of Entomology, One Shields Avenue, University of California, Davis, CA 95616, U.S.A.

Abstract: *Concerns about pollinator declines have grown in recent years, yet the ability to detect changes in abundance, taxonomic richness, and composition of pollinator communities is hampered severely by the lack of data over space and time. Citizen scientists may be able to extend the spatial and temporal extent of pollinator monitoring programs. We developed a citizen-science monitoring protocol in which we trained 13 citizen scientists to observe and classify floral visitors at the resolution of orders or super families (e.g., bee, wasp, fly) and at finer resolution within bees (*superfamily Apoidea) only. We evaluated the protocol by comparing data collected simultaneously at 17 sites by citizen scientists (observational data set) and by professionals (specimen-based data set). The sites differed with respect to the presence and age of hedgerows planted to improve habitat quality for pollinators. We found significant, positive correlations among the two data sets for higher level taxonomic composition, honey bee (*Apis mellifera*) abundance, non-*Apis bee abundance, bee richness, and bee community similarity. Results for both data sets also showed similar trends (or lack thereof) in these metrics among sites differing in the presence and age of hedgerows. Nevertheless, citizen scientists did not observe approximately half of the bee groups collected by professional scientists at the same sites. Thus, the utility of citizen-science observational data may be restricted to detection of community-level changes in abundance, richness, or similarity over space and time, and citizen-science observations may not reliably reflect the abundance or frequency of occurrence of specific pollinator species or groups.*

Keywords: Apoidea, arthropod, bee, ecological monitoring, floral visitor, pollinator monitoring, volunteer observer

Evaluación de la Calidad de Datos de Comunidades de Polinizadores Tomados por Ciudadanos-Científicos

Resumen: *En los últimos años ha incrementado la preocupación por la declinación de polinizadores; sin embargo, la habilidad para detectar cambios en la abundancia, riqueza taxonómica y composición de las comunidades de polinizadores está limitada por la carencia de datos espaciales y temporales. Los ciudadanos científicos pueden ser capaces de ampliar la extensión espacial y temporal de los programas de monitoreo de polinización. Desarrollamos un protocolo de monitoreo para ciudadanos científicos en el que capacitamos a 13 ciudadanos científicos para observar y clasificar a los visitantes florales a nivel de órdenes o superfamilias (e. g., abeja, avispa, mosca) y a nivel más fino solo en el caso de abejas (Superfamilia Apoidea). Evaluamos el protocolo mediante la comparación de datos recolectados simultáneamente en 17 sitios por ciudadanos científicos (conjunto de datos de observación) y por profesionales (conjunto de datos basados en especímenes). Los sitios difirieron con respecto a la presencia y edad de cercos vivos sembrados para mejorar la calidad del hábitat para los polinizadores. Encontramos correlaciones positivas, significativas, entre los dos conjuntos de datos para la composición taxonómica de nivel superior, la abundancia de abejas (Apis mellifera), la abundancia de abejas distintas a Apis, la riqueza de abejas, y la similitud*

de las comunidades de abejas. Los resultados para ambos conjuntos de datos también mostraron tendencias similares (o la ausencia de las mismas) en esos parámetros entre sitios diferentes por la presencia y edad de los cercos vivos. Sin embargo, los ciudadanos científicos no observaron cerca de la mitad de los grupos de abejas recolectados por científicos profesionales en los mismos sitios. Por lo tanto, la utilidad de los datos de observación generados por ciudadanos científicos pueden estar restringidos a la detección de cambios en la abundancia, riqueza o similitud en el tiempo y espacio, y las observaciones de ciudadanos científicos pueden no reflejar confiablemente la abundancia o frecuencia de ocurrencia determinadas especies o grupos de polinizadores.

Introduction

Roughly 90% of the world's flowering plant species depend on animal pollinators to reproduce (Bawa 1990), and lack of pollination critically limits reproduction of many plant populations (Ashman et al. 2004; Aguilar et al. 2006). Therefore, recent studies demonstrating declines in important pollinator groups (e.g., Biesmeijer et al. 2006; Grixti et al. 2009) are enhancing decades-old concerns (Allen-Wardell et al. 1998) about the functional consequences of losses, declines, and alterations in pollinator communities (Potts et al. 2010). Both national and international initiatives have noted the lack of information on changes in pollinator communities and called for the establishment of regional, national, and global pollinator monitoring programs (International Pollinators Initiative 1999; National Research Council 2007), including the possible incorporation of volunteers or citizen scientists to extend the temporal and spatial extent of survey efforts (National Research Council 2007). The occurrence of some major pollinator groups, such as bees, is notoriously patchy over space and time (Roubik 2001; Williams et al. 2001), potentially necessitating a level of survey effort that would quickly become too costly to implement. Citizen-science programs have yielded data of sufficient quality for use in the conservation and management of various taxa (Turner 2003; Vance et al. 2003; Altizer et al. 2004), for example, monarch butterflies (*Danaus plexippus*) (Howard & Davis 2009).

Data collected by citizen scientists now form the basis of many extensive ecological monitoring programs in Europe (Schmeller et al. 2009) and North America (e.g., Link & Sauer 1998). Citizen scientists are involved in many different taxon-based monitoring efforts (e.g., North American Amphibian Monitoring Program [http://www.pwrc.usgs.gov/naamp/]) and activities (e.g., BioBlitzes [http://www.pwrc.usgs.gov/blitz/]), including monitoring of pollinator communities (e.g., bumble bees [http://beespotter.mste.uiuc.ed], and sunflower visitors [http://greatpollinatorproject.org/index.html]). Three benefits of integrating citizen scientists into ecological monitoring networks are extension of spatial and temporal sampling effort, reduction of costs, and educational and recreational benefits for citizens

(Bonney et al. 2009). In Europe volunteers spend 148,000 person days/year for an imputed annual cost savings of €4 million (Schmeller et al. 2009). Concerns exist, however, that data quality will be poor and that rare species will not be detected (Genet & Sargent 2003).

The quality of citizen-science data has been assessed in relatively few studies. For taxon-based monitoring schemes, such assessments have been conducted for citizen-science data collected on mammals (Newman et al. 2003), freshwater invertebrates (Fore et al. 2001), terrestrial invertebrates (Lovell et al. 2009), coral reef fishes (Darwall & Dulvy 1996), seahorses (Goffredo et al. 2004), invasive crab species (Delaney et al. 2008), and birds (Altizer et al. 2004). Although volunteer monitoring of vertebrates is frequently observational, via direct census (Darwall & Dulvy 1996; Turner 2003; Goffredo et al. 2004) or animal signs (Newman et al. 2003), volunteer monitoring schemes for invertebrates typically involve specimen collection.

Assessments of invertebrate monitoring have compared data collected by professionals and citizen scientists at multiple stages of the monitoring process (i.e., collection, specimen processing, and identification or classification) (Fore et al. 2001; Nerbonne & Vondracek 2003; Lovell et al. 2009). In general, results of these studies show that citizen scientists can be trained quickly to collect samples that are highly similar in species (or higher taxon) richness and composition to those collected by professionals and to classify collected specimens of invertebrates with relatively high accuracy. For example, data from samples of benthic stream invertebrates collected by citizen scientists documented the same trends in invertebrate community response to urban development as samples collected by professionals (Fore et al. 2001), and citizen scientists achieved 80–95% accuracy in distinguishing among three crab species (Delaney et al. 2008). Nevertheless, citizen scientists often record fewer taxa than professionals (Fore et al. 2001). Citizen scientists may not detect smaller, less apparent, or highly mobile organisms and may lump taxa or misclassify certain taxonomic groups (Fore et al. 2001; Nerbonne & Vondracek 2003).

In contrast to prior work on invertebrate monitoring, we developed and evaluated an observational method for

citizen-science monitoring of abundance and taxonomic richness of pollinators in accordance with the U.S. National Research Council (2007) recommendation. A U.S. national citizen-science monitoring system would need to be observational, not only to minimize the possible effect on the pollinator fauna, but also to reduce the cost of processing and identifying specimens (National Research Council 2007). Indeed, it is unlikely that the capacity to identify a large volume of pollinator specimens exists. Taxonomists for the major groups (Hymenoptera [especially Apoidea], Diptera, and Lepidoptera) are already overwhelmed by their existing workloads (International Pollinators Initiative 1999). To assess the effectiveness of the observational monitoring protocol for pollinators, we compared it with methods based on identification of specimens (hereafter specimen-based methods) because specimen-based methods are typically used in scientific surveys of pollinator abundance and taxonomic richness (Brosi et al. 2007; Winfree et al. 2007; Westphal et al. 2008) and because it is necessary to examine specimens in the lab to identify individuals to the species level. We did not compare the accuracy of observational data collected by citizen versus professional scientists, although such data would be useful in determining sources of discrepancy between the two methods of data collection. Instead, we compared citizen-science data with the most likely professional alternative.

We investigated whether observational data on the abundance of floral visitors collected by citizen scientists correlate with specimen-based data at higher taxonomic levels (e.g., order, superfamily). Within bees (superfamily Apoidea) only, we investigated the data at a finer taxonomic resolution to determine whether observational data collected by citizen scientists correlate with specimen-based data for abundance, taxonomic richness, and composition of bee groups among sites. We also sought to determine whether the patterns in bee communities revealed by citizen-science data were the same as those revealed in specimen-based data among sites that differed with respect to presence and age of hedgerows.

Methods

Study Area

We conducted our study in Yolo County, a largely agricultural region in California's Central Valley. Research sites were located at the edges of farm fields along ditches, canals, or sloughs. At each of 17 sites, data were collected along a 350 m transect approximately 3 m wide. These sites are part of a long-term project in which the effects of planting hedgerows of native shrubs, grasses, and forbs (Bugg et al. 1998) on pollinator communities are being investigated (Menz et al. 2010). Two sites had mature hedgerow plantings (>10 years old), 5 sites had

new hedgerow plantings (<2 years old), and 10 sites had no hedgerow plantings. Hedgerows were typically 350–500 m long and about 3 m wide. Adjacent agricultural fields consisted of row-crops, alfalfa hay, grapes, or fruit trees. To ensure spatial independence, sites were located at least 1 km apart.

Training

In June 2008, 13 citizen scientists received 2 days of training (5 h each day). The citizen scientists were heterogeneous by gender (five males, eight females), education (six with undergraduate degrees, four with graduate degrees, two with "some college classes," one education not reported), and scientific experience. Four of the citizen scientists had no scientific experience, but the remainder had some laboratory or field experience, including four with experience identifying terrestrial invertebrates. None of the trainees had previous experience monitoring floral visitors. Trainees attended presentations on identifying and monitoring floral visitors, studied pinned specimens, and practiced observing and identifying visitors at flowers. During the practicum, each trainee, accompanied by a professional scientist with expertise in monitoring floral visitors, visited an organic farm with a hedgerow planting and diverse pollinator community (C.K., personal observation). On day 1 citizen scientists learned to distinguish among higher level taxonomic categories (e.g., bee, fly, wasp, other) of floral visitors. On day 2 they learned to distinguish among groups and subgroups of bees according to a hierarchical scheme (Fig. 1). Some of these groups or subgroups corresponded to a single species (e.g., yellow faced bumble bee = *Bombus vosnesenskii*), whereas others corresponded to family (e.g., hairy belly bee = Megachilidae) (Table 1). Key identification characteristics included size, shape, striping pattern, color, placement of pollen, hair distribution, and flight pattern (Ullmann et al. 2010).

Citizen scientists were trained to classify their observations to the highest possible level of resolution within the hierarchy (Fig 1). Thus, a citizen scientist could identify, with progressively higher taxonomic resolution, a yellow-faced bumble bee (*B. vosnesenskii*) as "bee," "non-*Apis*" bee (i.e., to differentiate from the honey bee, which is the dominant non-native species in the study area), "bumble bee," or "yellow-faced bumble bee."

Each citizen scientist received continuous feedback during one-on-one training sessions in the field from an accompanying professional scientist, who assessed the accuracy of each observation made by the citizen scientist. Professionals also assured consistency in monitoring methods, which required identifying only insects landing on the reproductive parts of the flower for ≥1 s and identifying those insects to the most highly resolved level in the hierarchy at which citizen scientists felt confident. Although not all floral visitors are pollinators and the
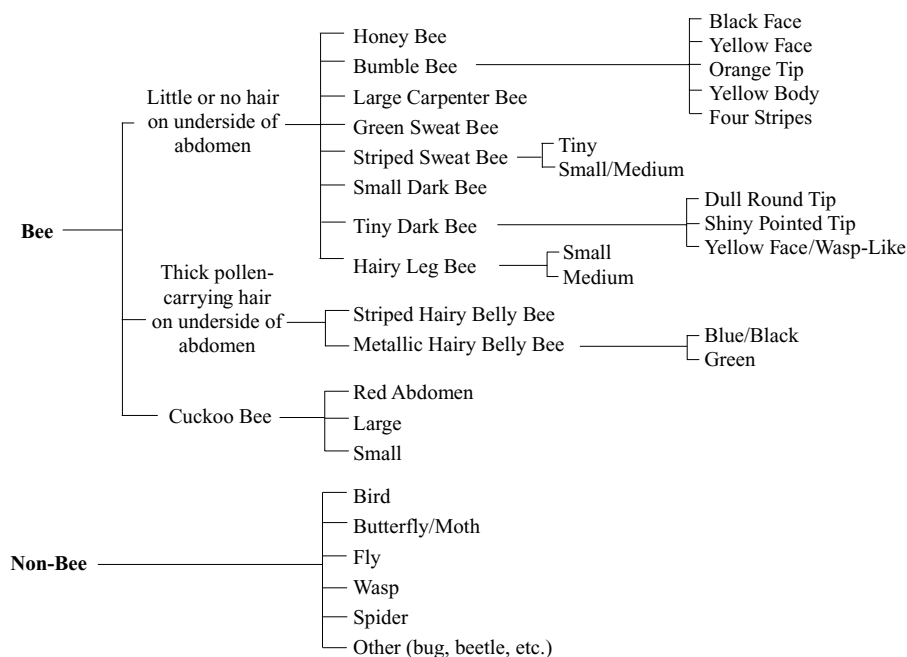
*Figure 1. Groups of flower visitors and groups and subgroups within bees used in data collection by citizen scientists. Non-*Apis *bees are all bee groups except for "honey bee" (*Apis mellifera*) and are primarily native species.*

effectiveness of different floral visitor species as pollinators of different plant species varies extensively from species to species (Herrera 2000; Kremen et al. 2002), meta-analyses show that abundance or visitation rate is a more important determinant than species-specific per-visit effectiveness in determining pollination function (Vázquez et al. 2005).

The nine professionals varied in years of experience (eight of the nine had > 3 years of experience in field sampling and curation of bee specimens, mean = 10.7; median = 5; 7 also had formal training in bee identification). During the 2 days of training, each trainee worked individually in the field with 2–3 different professionals. If trainees received training from one of the two professionals lacking formal training, they also received training from a more experienced professional.

Data Collection

At each site we randomly selected a pair of citizen scientists to collect observational data while a single professional (four of the nine from those described above, mean = 4.25 years experience; median = 3 years) collected specimens at the same site. Each of the 17 sites was sampled on a single date between 18–25 June 2008 (hereafter, first sample round), and 8 sites were resampled once between 4–12 August 2008 (second sample round). In June each citizen scientist sampled an average of 2.6 sites and in August an average of 3.2 sites. Sites were sampled between the hours of 0900 and 1400 when the temperature was ≥ 21 °C and wind was ≤ 2.5 m/s and when it was sunny, partly cloudy, or thinly overcast.

At each site each citizen scientist collected observational data on floral-visitor abundance and identity for

30 min while their citizen-scientist partner recorded the information on a data sheet, for a total of 1 h of observational data/day. Citizen scientists made observations while slowly walking along the 350 m transect, stopping briefly at all flower patches. They counted all floral visitors unless they knew an observation was of the same individual. Simultaneously, the professional collected specimens by netting floral visitors and by trapping insects in colored bowls (pan traps) (Westphal et al. 2008). Professionals collected floral visitors while walking along the transect at a steady pace during an effective 1 h of sampling (stopping the clock during specimen handling). Professional and citizen scientists started sampling at opposite ends of the transect to avoid interference with each other. Prior to beginning transect walks (0800 or when temperature reached 18 °C), the professional placed 21, 177 mL (6 ounce) pan traps (Solo Cup Company, Highfield Park, Illinois) painted white, yellow, or fluorescent blue (seven of each color) along the transect every 16 m, alternating each color regularly. Traps were collected after 5 h. All bee specimens were pinned, labeled, and identified to species by R.W.T.

Data Coding

Following specimen identification, we also coded each specimen record in the database by the most finely resolved, matching group or subgroup names in our hierarchical classification scheme (Fig. 1, e.g., a specimen of *Ceratina nanula* would be coded as a tiny dark bee, shield tip, Table 1). This allowed us to compare citizen-scientists' observations with specimens based either on the species identification (Latin binomial) of each specimen or on the group or

Table 1. Bees commonly found in Yolo County, California, and groups and subgroups used by citizen scientists to identify bee visitors to flowers.

Group[a]	Family	Species
Bumble bee		
black face bumble bee	Apidae	*Bombus californicus*
four stripes bumble bee	Apidae	*Bombus melanopygus*
yellow body bumble bee	Apidae	*Bombus sonorus*
yellow face bumble bee	Apidae	*Bombus vosnesenskii*
Large carpenter bee	Apidae	*Xylocopa* spp.
Cuckoo bee		
small	Apidae	*Nomada* spp.
	Apidae	*Triepeolus* spp.
	Apidae	*Triepeolus* sp. nr. *lestes*
large	Megachilidae	*Coelioxys octodentata*
Hairy belly bee[b]		
Striped		
tiny	Megachilidae	*Ashmeadiella aridula astragali*
medium/large	Megachilidae	*Megachile apicalis*[c]
	Megachilidae	*Megachile brevis onobrychidis*
	Megachilidae	*Megachile coquilletti*
	Megachilidae	*Megachile frugalis*
	Megachilidae	*Megachile gentilis*
	Megachilidae	*Megachile rotundata*[c]
Metallic		
blue	Megachilidae	*Osmia* spp.
green	Megachilidae	*Osmia (Chenosmia)* spp.
Hairy leg bee		
small	Apidae	*Diadasia consosciata*
medium	Apidae	*Anthophora urbana*
	Apidae	*Diadasia enavata*
	Apidae	*Melissodes agilis*
	Apidae	*Melissodes lupina*
	Apidae	*Melissodes robustior*
	Apidae	*Melissodes stearnsi*
	Apidae	*Melissodes tepida timberlakei*
	Apidae	*Peponapis pruinosa*
large	Apidae	*Svastra obliqua expurgata*
Honey bee	Apidae	*Apis mellifera*[c]
Class		
Striped sweat bee		
small	Halictidae	*Halictus tripartitus*
medium	Halictidae	*Halictus ligatus*
Tiny dark bee		
dull round tip	Halictidae	*Lasioglossum incompletum*
	Halictidae	*Lasioglossum tegulariforme*
	Halictidae	*Lasioglossum (Dialictus)* spp.
shield tip	Apidae	*Ceratina acantha*
	Apidae	*Ceratina arizonensis*
	Apidae	*Ceratina dallatorreana*[c]
	Apidae	*Ceratina nanula*
yellow face and wasp like	Colletidae	*Hylaeus leptocephalus*
	Colletidae	*Hylaeus mesillae*
	Colletidae	*Hylaeus rudbeckiae*

[a]*Similar designations have been used to collect observational data on crop pollinators (e.g., Kremen et al. 2002). These groups and their family-level designations could likely be used in other geographic areas (see http://www.xerces.org/download/pdf/PA_Xerces%20Guide.pdf), but different species would occur within these categories.*
[b]*For the purpose of analysis the tiny striped hairy belly bee class was combined with the small/medium striped hairy belly bee class.*
[c]*Introduced species.*

subgroup classification of the specimen (corresponding to the taxonomic resolution of the observational data).

Citizen scientists sometimes identified individual bees at a less-resolved level in the hierarchy (Fig 1, e.g., striped sweat bee) and sometimes at a more-resolved level (small striped sweat bee). To calculate taxonomic richness from observations, the smallest number of independent bee groups was counted per site per sample date. For example, if both a less-resolved group (striped sweat bee) and

a more-resolved subgroup within it (small, striped sweat bee) were noted as present by the citizen scientist at a site on a particular date (hereafter site date), then only one group was added to the count of taxonomic richness.

To calculate community dissimilarity (e.g., relative Sorensen for use in Mantel testing), we removed uninformative data (e.g., small native bee) from the data set ($n = 17$ records). Then, for groups with < 30 total records, we combined the subgroups within these groups, to avoid creating many pseudo-rare groups. Specifically, we combined medium striped sweat bee ($n = 26$) with medium sweat bee ($n = 1$); medium ($n = 1$) and small striped hairy belly bee ($n = 5$) with striped hairy belly bee ($n = 6$); and medium hairy leg bee ($n = 8$) with hairy leg bee ($n = 2$).

Data Analyses

First, we compared the observational data with the netting data only because the netting method was the best analogue to the observational method (Westphal et al. 2008). Then we compared the observational data with the entire specimen-based data set (netting plus pan trapping). Pan trapping is an efficient passive sampling method that is free of observer or investigator bias (Westphal et al. 2008; but see Winfree et al. 2007). Many pollinator surveys typically collect both netting and pan-trapping data (e.g., Brosi et al. 2007). For this reason it was also important for us to know whether observational data collected by citizen scientists showed the same trends as data collected through the more complete and typical sampling scheme of netting and pan trapping. Prior to comparisons of observations with specimen-based data, we examined correlations and Bray-Curtis dissimilarity coefficients between netting and pan-trapping data.

We examined correlations between observational and specimen-based data for abundances within higher level taxonomic categories (Fig 1). Data on higher level taxonomic categories were lumped across sites and dates because there were too few individuals per site date for many of these taxa.

Within the higher taxon bees (Apoidea), we examined correlations between the observational and specimen-based data for *Apis mellifera* abundance (hereafter *Apis*), non-*Apis* abundance, taxonomic richness (number of bee groups or subgroups for observations compared with species richness or group or subgroup richness for specimens), and proportions of non-*Apis* individuals on each site date. Most non-*Apis* bee species in our sample were native, but a few naturalized species occurred at low abundances (*Megachile apicalis*, *M. rotundata*, and *Ceratina dallatorreana*, together representing 0.9% of the sample). Site date was our unit of replication because we were interested in differences (spatial or temporal) in bee community attributes between site dates and whether

these differences could be detected equally well by analysis of either observational or specimen-based data (Lovell et al. 2009). Prior to analysis, we examined whether sample period (first or second round) influenced the specimen data, but we found no statistically significant differences ($p > 0.05$) in *Apis* abundance, non-*Apis* abundance, proportion of non-*Apis* individuals, or species and bee group richness for data collected either by netting or by netting and pan trapping.

We used Mantel tests to determine whether patterns of community dissimilarity between site dates were correlated for the observational versus the specimen-based data sets. We calculated distance matrices for each data set with the relative Sorensen distance measure to obtain distances between all pairs of site dates. The Mantel test randomly permutes one of the two distance matrices being compared (here, observational data vs. either netting data or netting and pan-trapping data) and computes correlations with the unpermuted matrix to develop an empirical null distribution of correlation coefficients against which to test the correlation between the two original matrices, thereby avoiding partial dependence within the matrix (McCune & Grace 2002). We used the standardized Mantel statistic r and evaluated its statistical significance with Monte Carlo randomization (1000 trials) in PC ORD (version 4.0, MJM Softward Design, Gleneden Beach, Oregon).

To examine whether analysis of observational data revealed differences in bee communities among hedgerow classes (mature, new, none), we used generalized linear models (GLM). We used only data from the first sampling period because mature hedgerow sites were not sampled during the second sample period. First, we tested the specimen-based data to determine whether hedgerow class affected abundance or richness. Then we conducted parallel tests with the observational data. The goal of this analysis was not to test the effect of hedgerow class on pollinator communities, but to determine whether analysis of citizen-scientist observational data yielded similar inferences about these potential effects as specimen-based data.

Using GLM we investigated explicitly the effect of sample period on the accuracy of observations to determine whether accuracy changed in the second sample round. Changes in accuracy might occur due to greater experience in the second round (positive change expected) or to greater time elapsed since the training (negative change expected). Bee groups that were detected by both citizen scientists and professionals at a site date were those groups known to be detected correctly by citizen scientists. In contrast, groups detected by citizen scientists but not occurring in the specimen data set could be either false positives (commissions by citizen scientists) or the result of sampling error (omissions) by professionals. We therefore used proportion of groups detected by both observations and specimens relative to

the total number of groups detected in both data sets as the response variable for the effect of sample period on accuracy.

To identify sources of discrepancy between observational and specimen-based data sets, we categorized the data at each site date by bee group as follows: correct, recorded in both data sets; undetected, recorded in specimen-based data set only; observational data only, detected by observations but specimens not collected. Observations that were not fully resolved within our hierarchy (e.g., striped sweat bee, Fig 1) were queried against the corresponding fully resolved group in the specimen-based data (e.g., small striped sweat bee or medium striped sweat bee), but only if the fully resolved group had not also been observed at the site date (as for species richness). We categorized observational data as correct but not fully resolved if specimens in the corresponding fully resolved group had been collected at the site or as observed but not detected in specimen data (as above) if no specimens were collected. For each bee group, we calculated the proportion of discrepancies occurring between the citizen-science data and the specimen-based data as (undetected + observational data only)/(correct + undetected + observational data only).

We used Spearmans' rank test for correlations among variables that were non-normally distributed because many variables did not respond to common transformations, and Pearson's correlation for normally distributed variables (species and group richness). For GLM we used analysis of variance for normally distributed variables and negative binomial regression for overdispersed variables (all abundance variables). For statistically significant models, we assessed differences among groups with the Wald test. We used JMP (version7.0, SAS Institute, Cary, North Carolina) to perform correlations and Stata (version 7.0, Stata Corporation, College Station, Texas) to conduct GLM.

Results

Across the 25 site-date combinations, professionals collected 2119 insects in nets and 4548 insects in nets plus pan traps, whereas citizen scientists observed 3763 insects. There was no significant correlation between the netting versus the pan-trapping data sets by site date for *Apis* abundance, non-*Apis* abundance, proportion of non-*Apis* individuals, species richness, or bee group richness ($p \gg 0.05$). The two data sets differed in species composition, with mean dissimilarity (Bray-Curtis Index) of 0.69 (SD 0.15; $n = 25$ site dates; 0, complete similarity; 1, complete dissimilarity). Mean dissimilarity based on bee group composition was similarly high (0.67 [0.16]). Of the 43 species detected across all sites, 23 (56%) were detected by both methods. Thirty-six species (84%) were caught in nets and 31 (72%) in pan traps.
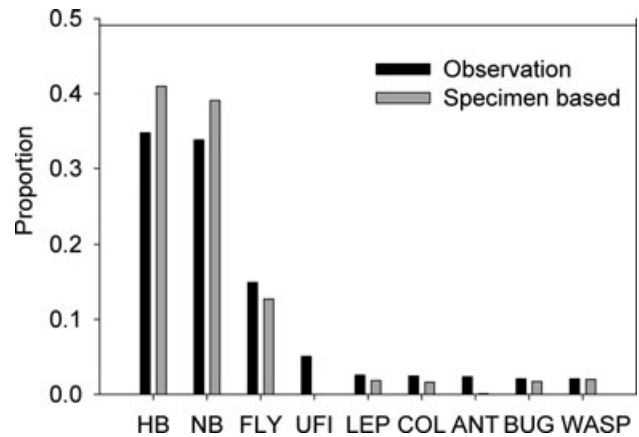

Figure 2. Proportion of observations and specimens (netting data) identified as honey bees (HB), non-Apis bees (NB), flies (FLY), unidentified flying insects (UFI), butterflies or moths (LEP), beetles (COL), ants (ANT), bugs (BUG), and wasps (WASP). Observation data were significantly positively correlated with specimen data (Spearman's $\rho = 0.68$, p < 0.05).

Comparisons between Observational and Specimen-Based Data

Higher level taxonomic groups were recorded in similar proportions in observational and specimen-based data sets (sites and dates combined, Fig. 2, netting: $\rho = 0.68$, $p < 0.05$; netting and pan trapping: $\rho = 0.79$, $p = 0.01$). *Apis* and non-*Apis* bees were the most abundant insects, but collections also included wasps, flies, and other insects. Forty-eight species of non-*Apis* bees in four families (Apidae, Colletidae, Halictidae, Megachilidae) were identified from the collected specimens.

Comparisons of observations with netting data for *Apis* abundance, non-*Apis* abundance, taxonomic richness (at the species and bee group level for specimen-based data), and proportion of non-*Apis* individuals were all significantly and positively correlated (Table 2, Fig. 3). Similar significant results were obtained for comparisons of observations with netting plus pan-trapping data (with smaller correlation coefficients, Table 2), except for taxonomic richness when compared at the species level for specimen-based data.

Community dissimilarity among sites was correlated significantly between observational and specimen-based data sets for both the bee group (observational) to bee group (specimen-based) comparison (netting: $r = 0.489$, $p = 0.001$; netting and pan trapping: $r = 0.310$, $p = 0.001$) and the bee group (observational) to species-level (specimen-based) comparison (netting: $r = 0.475$, $p = 0.001$; netting and pan trapping: $r = 0.324$, $p = 0.001$).

For specimen-based data, we detected no trend in non-*Apis* bee abundance among hedgerow classes (GLM, netting: $z = 0.79$, $p = 0.43$; netting plus pan trapping:

**Table 2. Correlations between observational and specimen-based data for various response variables[a].**

Comparison	Netting data	Netting and pan-trapping data
	Spearman's (ρ, p)	
Apis mellifera abundance	0.92, < 0.001	0.85, < 0.001
Non-Apis bee abundance	0.88, < 0.001	0.44, < 0.05
Proportion of non-Apis bees	0.59, < 0.005	0.60, < 0.005
	Pearson's (r, p)	
Group to group richness[a]	0.61, = 0.001	0.475, < 0.05
Group to species richness[b]	0.65, < 0.001	0.26, > 0.05

[a] *For specimen-based data, richness is calculated on the basis of bee groups.*
[b] *For specimen-based data, richness is calculated on the basis of species-level identifications.*

$z = -0.09$, $p = 0.93$). Species richness was significantly higher in sites with mature hedgerows than in sites with new hedgerows or no hedgerows (GLM, netting: $F_{2,14} = 6.27$, $p = 0.01$, Wald test mature≫new = none, $p < 0.01$; netting and pan trapping: $F_{2,14} = 7.11$, $p < 0.01$, Wald test mature≫new = none, $p < 0.01$). Similarly, the observational data showed no differences among hedgerow classes for non-*Apis* bee abundance (GLM: $z = 0.24$, $p = 0.81$), but showed a significantly higher richness in sites with mature hedgerows than in sites with new hedgerows or no hedgerows (GLM: $F_{2,14} = 8.07$, $p = 0.005$; Wald test: mature≫new = none, $p < 0.005$).

On average the citizen scientists detected 48% (SE 4) of the fully resolved bee groups collected by netting at the same sites and an additional 8% (1.6) of bee groups with partial taxonomic resolution (Fig 1 & Table 1, e.g., tiny dark bee instead of tiny dark bee / dull round tip). There was no significant difference between the first and second sample rounds in the proportion of groups recorded in both observational and specimen-based data sets at a site date (GLM ANOVA: $F_{1,24} = 0.47$, $p = 0.50$). On average, 1.8 bee groups (SE 0.2) were undetected in observations, but observations included 1.2 bee groups per site (0.2) not sampled by netting. Similar results were obtained when comparing citizen-science detection rates with netting and pan-trapping specimens, but the fully resolved detection rates dropped to 41% (SE 3), with 3.2 bee groups per site (0.2) undetected and reporting of 0.96 bee groups per site (0.14) not sampled by netting or pan trapping. Across all bee groups we found a higher proportion of discrepancies between the observational and specimen-based data for bee groups that had lower relative abundance (netting: ρ = –0.69, p <0.005; netting and pan trapping: ρ = –0.59, $p = 0.01$) or frequency of
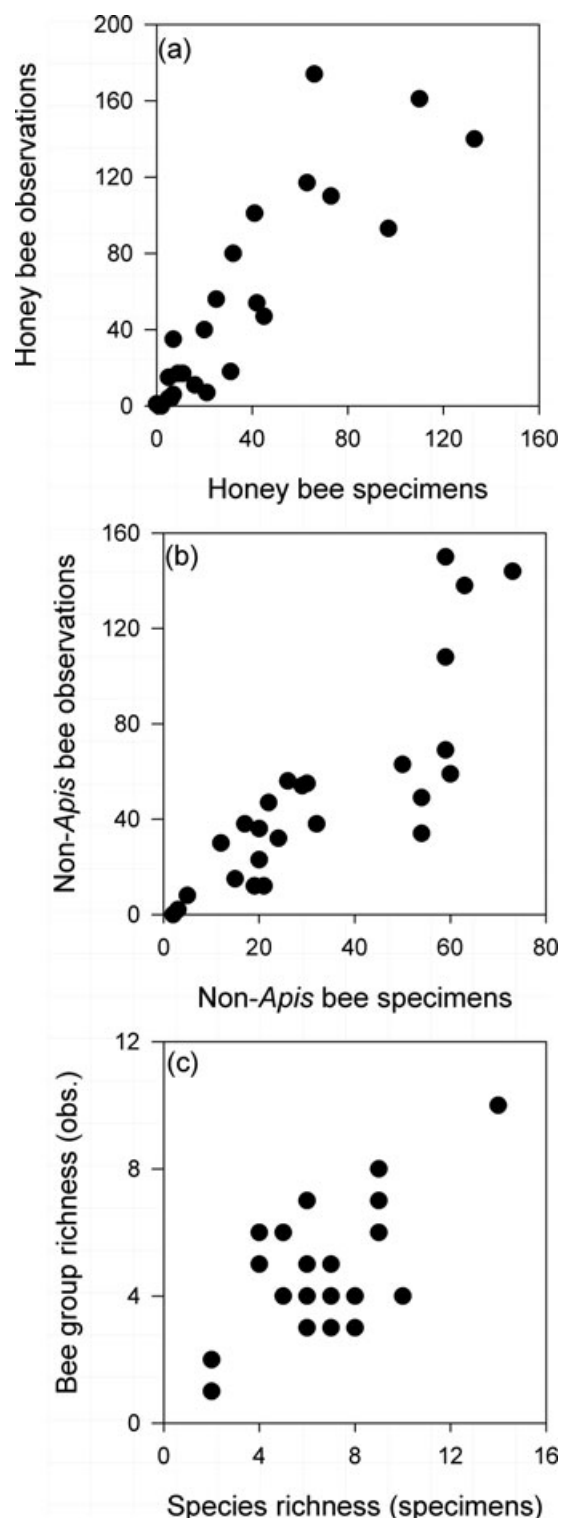
Figure 3. Among-site trends in bee abundance and taxonomic richness from analysis of observations and specimens (netting data): (a) Apis *abundance (Spearman's ρ = 0.92, p < 0.001), (b) non-*Apis *abundance (Spearman's ρ = 0.88, p < 0.001), and (c) richness of bee groups (observations) versus species richness of specimen-based data (Pearson's* r = 0.65, p < 0.001) (each point is a site date).

occurrence (netting: $\rho = 0.58, p = 0.01$; netting and pan trapping: $\rho = 0.46, p = 0.06$ [Supporting Information]).

Discussion

Our results suggest that citizen scientists with modest training can collect useful observational data for detecting spatial and temporal differences in pollinator community attributes. Observational data were significantly and positively correlated to the netting or netting and pan-trapping data for all but one of the response variables we measured. The observational data mirrored the specimen-based data when we compared sites with different hedgerow classes. None of the data sets showed a significant change in non-*Apis* bee abundance among hedgerow classes, whereas all of the data sets showed a significantly higher richness of bee species at the mature hedgerow sites. Our results are consistent with results of other similar studies (e.g., Darwall & Dulvy 1996; Fore et al. 2001) which suggest that useful data for monitoring changes in communities over space and time can be collected by nonprofessionals (National Research Council 2007). These techniques might be used, for example, by natural resource managers, growers, field researchers, or backyard gardeners to monitor pollinator communities' responses to alternative land-management practices.

Our sampling design confounded methodological differences (observation vs. sampling) with experience level (nonprofessional vs. professional), which limits the inferences that can be made regarding the causes of discrepancies between the two data sources. Discrepancies between observed and specimen-based data were greatest for the bee-identification groups that were low in abundance, occurred infrequently, and were patchily distributed or hard to identify (Supporting Information). These differences may have resulted from misclassifications by citizen scientists or from true differences among the samples. We suggest that future assessments of observational data for invertebrate communities include observational data from both citizen and professional scientists, in addition to collection of specimens, in order to distinguish methodological differences from experience level.

If discrepancies between the data sets were chiefly due to misclassifications, then additional training may be required so that citizen scientists are sufficiently familiar with all bee groups, particularly the rarer ones. Desirable modifications to the training protocol may include reducing subjectivity in the classifications (as suggested by Lovell et al. 2009) by better defining the size classes and practicing sizing objects bracketing the range of bee sizes, independent of observing and classifying bees in the field (Darwall & Dulvy 1996).

True differences may also exist between the observational and the specimen-based data. Bees have considerable spatial and temporal turnover even over small extents (Roubik 2001; Williams et al. 2001); citizen scientists working at one end of the transect may have observed individuals in a given bee group that only occurred in that part of the transect at that moment in time and thus were not collected by the professionals, who worked from the opposite end of the transect. These sampling differences could be reduced by increasing the duration of sampling for both observations and collection of specimens. The difference in methods may also explain why we obtained lower correlations than previous researchers who compared data collected by citizens and professionals in which both groups used identical collection methods (e.g., Fore et al. 2001). In addition, observations of bee species richness and abundance by professionals are known to be less efficient in the proportion of species detected relative to other specimen-based techniques (Westphal et al. 2008).

We found that the observational data were more strongly correlated with netting data, than with netting plus pan-trapping data. The major differences between observing floral visitors as a sampling technique and collecting visitors at flowers are that collection occurs over a longer period than observation to achieve the same sample effort, may affect adjacent visitors, samples without replacement, and has different sample biases than observation. Despite these differences, we found strong correlations between the observational data and the netting data. Furthermore, we found the netting and pan-trapping data sets were uncorrelated for all metrics and thus complementary in their assessment of community composition (Potts et al. 2005, but see Westphal et al. 2008). That we nonetheless found positive correlations between the observational data and the netting and pan-trapping data suggest that the observational data collected by citizen scientists may be able to represent trends in abundance, richness, and composition of pollinator communities even though such data are less complete than data produced by typical collection-sampling schemes. Citizen-science observational monitoring of pollinator communities could allow the extent of monitoring programs to increase without the danger of overcollection and without overtaxing the scarce human resource available to identify collected material.

As in other studies, we found that although citizen scientists can readily be trained to collect data that allow detection of trends in community attributes, biases exist in the types of species or higher taxa they sample (Fore et al. 2001; Lovell et al. 2009). In our study citizen scientists sampled by observing flower visitors. Lovell et al. (2009) found that including two citizen scientists per professional was sufficient to remove sampling differences between data collected by citizen scientists and professionals. Thus, a potential correction for the

tendency of citizen scientists to undersample relative to professionals could be to increase the citizens' observer effort (Schmeller et al. 2009).

We found, as others have, that biases also exist in the types of species or higher taxa that citizen scientists are able to classify correctly (Fore et al. 2001; Nerbonne & Vondracek 2003). In general, the degree of bias in identification or classification accuracy varies depending on the difficulty of classifying groups or species within a given taxon and the taxonomic breadth of the monitoring scheme. For example, the citizen scientists in Delaney et al.'s (2008) study attained high levels of accuracy in a sampling scheme that was narrowly defined (three crab species), whereas in sampling of stream invertebrates, a much broader and more difficult task, citizen scientists correctly classified <30% of individuals to the family level (Nerbonne & Vondracek 2003). In our study, which included both taxonomic breadth and many small insects that are difficult to observe and classify, citizen scientists detected less than half of the bee groups that professional researchers obtained by collection.

Despite these discrepancies, trends in community-level attributes of abundance, richness, and composition derived from the citizen-science data presented similar across- and between-site trends to those obtained by professional scientists. Further refinements of our classification system and training protocol could improve the accuracy of observational pollinator monitoring. With modest adaptations of bee groups (dependent on the resident fauna), the classification system (Fig 1) could be made applicable for use in other locations in North America—one such adaptation is already in use in Pennsylvania (http://www.xerces.org/download/pdf/PA_Xerces%20Guide.pdf). A larger data set is needed to allow calibration of observational data with specimen-based data for use in a national pollinator-monitoring program as recommended by (National Research Council 2007).

Acknowledgments

Supporting information

Results of analysis of discrepancies between observational and specimen-based data are available as part of the online article (Appendix S1). The authors are responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

Literature Cited

Aguilar, R., L. Ashworth, L. Galetto, and M. A. Aizen. 2006. Plant reproductive susceptibility to habitat fragmentation: review and synthesis through a meta-analysis. Ecology Letters **9:**968–980.

Allen-Wardell, A. G., et al. 1998. The potential consequences of pollinator declines on the conservation of biodiversity and stability of food crop yields. Conservation Biology **12:**8–17.

Altizer, S., W. M. Hochachka, and A. A. Dhondt. 2004. Seasonal dynamics of mycoplasmal conjunctivitis in eastern North American house finches. Journal of Animal Ecology **73:**309–322.

Ashman, T.-L., et al. 2004. Pollen limitation of plant reproduction: ecological and evolutionary causes and consequences. Ecology **85:**2408–2421.

Bawa, K. S. 1990. Plant-pollinator interactions in tropical rain forests. Annual Review of Ecology and Systematics **21:**299–422.

Biesmeijer, J. C., et al. 2006. Parallel declines in pollinators and insect-pollinated plants in Britain and The Netherlands. Science **313:**351–354.

Bonney, R., C. B. Cooper, J. Dickinson, S. Kelling, T. Phillips, K. V. Rosenberg, and J. Shirk. 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. BioScience **59:**977–984.

Brosi, B. J., G. C. Daily, and P. R. Ehrlich. 2007. Bee community shifts with landscape context in a tropical countryside. Ecological Applications **17:**418–430.

Bugg, R. L., J. H. Anderson, C. D. Thomsen, and J. Chandler. 1998. Farmscaping in California: hedgerows, roadside plantings and wild plants for biointensive pest management. Pages 339–374 in C. H. Pickett, and R. L. Bugg, editors. Enhancing biological control: habitat management to promote natural enemies of agricultural pests. University of California Press, Berkeley.

Darwall, W. R. T., and N. K. Dulvy. 1996. An evaluation of the suitability of non-specialist volunteer researchers for coral reef fish surveys. Mafia Island, Tanzania—a case study. Biological Conservation **78:**223–231.

Delaney, D. G., C. D. Sperling, C. S. Adams, and B. Leung. 2008. Marine invasive species: validation of citizen science and implications for national monitoring networks. Biological Invasions **10:**117–128.

Fore, L. S., K. Paulsen, and K. O'Laughlin. 2001. Assessing the performance of volunteers in monitoring streams. Freshwater Biology **46:**109–123.

Genet, K. S., and L. G. Sargent. 2003. Evaluation of methods and data quality from a volunteer-based amphibian call survey. Wildlife Society Bulletin **31:**703–714.

Goffredo, S., C. Piccinetti, and F. Zaccanti. 2004. Volunteers in marine conservation monitoring: a study of the distribution of seahorses carried out in collaboration with recreational scuba divers. Conservation Biology **18:**1492–1503.

Grixti, J. C., L. T. Wong, S. A. Cameron, and C. Favret. 2009. Decline of bumble bees (*Bombus*) in the North American Midwest. Biological Conservation **142:**75–84.

Herrera, C. M. 2000. Flower-to-seedling consequences of different pollination regimes in an insect-pollinated shrub. Ecology **81:**15–29.

Howard, E., and A. K. Davis. 2009. The fall migration flyways of monarch butterflies in eastern North America revealed by citizen scientists. Journal of Insect Conservation **13:**279–286.

International Pollinators Initiative. 1999. The São Paulo declaration on pollinators. Page 66 in Report on the recommendations of the workshop on the conservation and sustainable use of pollinators in agriculture with emphasis on bees. Brazilian Ministry of the Environment, Brasilia.

Kremen, C., N. M. Williams, and R. W. Thorp. 2002. Crop pollination from native bees at risk from agricultural intensification. Proceedings of the National Academy of Sciences **99:**16812–16816.

Link, W. A., and J. R. Sauer. 1998. Estimating population change from count data: application to the North American Breeding Bird Survey. Ecological Applications **8:**258–268.

Lovell, S., M. Hamer, R. Slotow, and D. Herbert. 2009. An assessment of the use of volunteers for terrestrial invertebrate biodiversity surveys. Biodiversity and Conservation **18:**3295–3307.

McCune, B., and J. B. Grace 2002. Analysis of ecological communities. MjM Software Design, Gleneden Beach, Oregon.

Menz, M. H. M., R. D. Phillips, R. Winfree, C. Kremen, M. A. Aizen, S. D. Johnson, and K. W. Dixon. 2010. Reconnecting plants and pollinators: challenges in the restoration of pollination mutualisms. Trends in Plant Science DOI: 10.1016/j.tplants.2010.09.006.

National Research Council. 2007. Status of pollinators in North America. National Academy Press, Washington, D.C.

Nerbonne, J. F., and B. Vondracek. 2003. Volunteer macroinvertebrate monitoring: assessing training needs through examining error and bias in untrained volunteers. Journal of the North American Benthological Society **22:**152–163.

Newman, C., C. D. Buesching, and D. W. Macdonald. 2003. Validating mammal monitoring methods and assessing the performance of volunteers in wildlife conservation—"Sed quis custodiet ipsos custodies ?" Biological Conservation **113:**189–197.

Potts, S. G., J. C. Biesmeijer, C. Kremen, P. Neumann, O. Schweiger, and W. E. Kunin. 2010. Global pollinator declines: trends, impacts and drivers. Trends in Ecology and Evolution **25:**345–353.

Potts, S. G., P. G. Kevan, and J. W. Boone. 2005. Conservation in pollination: collecting, surveying and monitoring. Pages 401–434 in A. Dafni, and P. G. Kevan, editors. Pollination ecology: a practical approach. Enviroquest, Cambridge, Canada.

Roubik, D. 2001. Ups and downs in pollinator populations: when is there a decline? Conservation Ecology **5:**2 http://www.ecologyandsociety.org/vol5/iss1/art2/ (accessed February 2011).

Schmeller, D. S., et al. 2009. Advantages of volunteer-based biodiversity monitoring in Europe. Conservation Biology **23:**307–316.

Turner, W. R. 2003. Citywide biological monitoring as a tool for ecology and conservation in urban landscapes: the case of the Tucson Bird Count. Landscape and Urban Planning **65:**149–166.

Ullmann, K., M. Vaughan, C. Kremen, T. Shih, and M. Shepherd. 2010. California Pollinator Project-citizen scientist pollinator monitoring guide. The Xerces Society for Inverterbrate Conservation, Portland, Oregon. Available from http://www.xerces.org/wp-content/uploads/2010/06/CA_CSM_guide.pdf (accessed July 2010).

Vance, M. D., L. Fahrig, and C. H. Flather. 2003. Effect of reproductive rate on minimum habitat requirements of forest-breeding birds. Ecology **84:**2643–2653.

Vázquez, D. P., W. F. Morris, and P. Jordano. 2005. Interaction frequency as a surrogate for the total effect of animal mutualists on plants. Ecology Letters **8:**1088–1094.

Westphal, C., et al. 2008. Measuring bee diversity in different european habitats and biogeographical regions. Ecological Monographs **78:**653–671.

Williams, N. M., R. L. Minckley, and F. A. Silveira. 2001. Variation in native bee faunas and its implications for detecting community changes. Conservation Ecology **5:**7 http://www.ecologyandsociety.org/vol5/iss1/art7/ (accessed February 2011).

Winfree, R., T. Griswold, and C. Kremen. 2007. Effect of human disturbance on bee communities in a forested ecosystem. Conservation Biology **21:**213–223.