Ecological Applications, 20(8), 2010, pp. 2170–2187

Unite research with what citizens do for fun: "recreational monitoring" of marine biodiversity

STEFANO GOFFREDO,[1,3] FRANCESCO PENSA,[1] PATRIZIA NERI,[1] ANTONIO ORLANDI,[1] MARIA SCOLA GAGLIARDI,[1] ANGELA VELARDI,[1] CORRADO PICCINETTI,[1,2] AND FRANCESCO ZACCANTI[1]

[1] *Marine Science Group, Citizen Science Lab, Department of Evolutionary and Experimental Biology, Alma Mater Studiorum, University of Bologna, Via F. Selmi 3, I-40126 Bologna, Italy*
[2] *Laboratory of Fisheries and Marine Biology at Fano, Alma Mater Studiorum, University of Bologna, Viale Adriatico 1/N, I-61032 Fano (PU), Italy*

Abstract. Institutes often lack funds and manpower to perform large-scale biodiversity monitoring. Citizens can be involved, contributing to the collection of data, thus decreasing costs. Underwater research requires specialist skills and SCUBA certification, and it can be difficult to involve volunteers. The aim of this study was to involve large numbers of recreational divers in marine biodiversity monitoring for increasing the environmental education of the public and collecting data on the status of marine biodiversity. Here we show that thousands of recreational divers can be enrolled in a short time. Using specially formulated questionnaires, nonspecialist volunteers reported the presence of 61 marine taxa encountered during recreational dives, performed as regular sport dives. Validation trials were carried out to assess the accuracy and consistency of volunteer-recorded data, and these were compared to reference data collected by an experienced researcher. In the majority of trials (76%) volunteers performed with an accuracy and consistency of 50–80%, comparable to the performance of conservation volunteer divers on precise transects in other projects. The recruitment of recreational divers involved the main diving and tour operators in Italy, a popular scientific magazine, and mass media. During the four-year study, 3825 divers completed 18 757 questionnaires, corresponding to 13 539 diving hours. The volunteer-sightings-based index showed that in the monitored area the biodiversity status did not change significantly within the project time scale, but there was a significant negative correlation with latitude, suggesting improved quality in the southernmost areas. This trend could be related to the presence of stressors in the northern areas and has been supported by investigations performed by the Italian Ministry of the Environment. The greatest limitation with using volunteers to collect data was the uneven spatial distribution of samples. The benefits were the considerable amounts of data collected over short time periods and at low costs. The successful development of citizen-based monitoring programs requires open-mindedness in the academic community; advantages of citizen involvement in research are not only adding large data sets to the ecological knowledge base but also aiding in the environmental education of the public.

Key words: biodiversity; citizen science; education; environmental monitoring; Mediterranean Sea; SCUBA divers; volunteers in research.

INTRODUCTION

Preserving biodiversity and the benefits it provides to society is a basic need for mankind (Balmford et al. 2005). The identification and quantification of threats enable managers to take effective measures. While broad conservation efforts require the implementation of global monitoring programs to build up-to-date databases, government agencies are often under-funded, and many cannot afford large-scale monitoring (Sharpe and Conrad 2006). Paradoxically, this decline in ecological monitoring over the second half of the 20th century has coincided with the huge increase in concern for biodiversity and the environment (Secord 1996). Economic constraints on data collection in some cases can be overcome by using the skills of nonspecialist volunteer researchers: the "citizen scientists" (Darwall and Dulvy 1996, Fore et al. 2001, Bhattacharjee 2005, Bell 2007, Greenwood 2007, Cohn 2008).

Citizen scientists are typically people who care about the wild, feel at home in nature, want to feel like they are making a difference while exploring new places, seek an experience where they help solve environmental problems, and have some awareness of the scientific process learning new things about nature (Gilmour and Saunders 1995, Ryan et al. 2001, Bruyere and Rappe 2007, Cohn 2008). They are attracted by the opportunity for cultural immersion, the chance to gain research

Manuscript received 25 August 2009; accepted 18 January 2010; final version received 11 February 2010. Corresponding Editor: P. K. Dayton.

[3] E-mail: stefano.goffredo@marinesciencegroup.org

experience, and the camaraderie that occurs on volunteer projects (Campbell and Smith 2006).

Citizen science contributed to the conservation of various organisms, adding information about their population structure, distribution, and behavior, and resource managers have taken advantage of volunteer networks (Darwall and Dulvy 1996, Fore et al. 2001, Goffredo et al. 2004, Bell 2007, Delaney et al. 2008). The United Nations Environment Program now emphasizes public involvement in environmental monitoring and management (Sharpe and Conrad 2006). The advantages of using such nonspecialist volunteers include the provision of manpower sufficient to conduct extensive surveys, providing simultaneous spatial coverage and placing the investigation in its local context; large financial savings through the provision of free labor and fund raising; an increase in the level of public awareness of ecological problems through active participation in ecological survey work; and the provision of a simple, low-cost survey program that can be continued in the long term using local expertise and financing (Stokes et al. 1990, Darwall and Dulvy 1996, Goffredo et al. 2004, Sheil and Lawrence 2004, Greenwood 2007). This is especially important since permanent monitoring increases the chance of early detection of biological invasions, and offers the greatest likelihood for their eradication (Myers et al. 2000, Lodge et al. 2006, Delaney et al. 2008).

The reliability and relevance of data generated by nonspecialist volunteers are held with some skepticism by the scientific community (Darwall and Dulvy 1996, Foster-Smith and Evans 2003), and despite the advantages raised above some seem reluctant to accept citizen science. The use of nonspecialist volunteers is often criticized on the grounds that the information collected will be unreliable as a result of either insufficient training or lack of consistency from using large numbers of observers (Darwall and Dulvy 1996). The potential of citizen science needs evaluation and its challenges need to be addressed since outright disregard means that valuable opportunities are being missed (Douglas and Lawrence 2004). Acceptance of citizen science by the scientific community would allow widespread nonspecialist participation in monitoring, and thereby greatly increase our ecological understanding by creating large spatial and temporal data sets.

For terrestrial environments, a range of successful ecological projects are based on the active involvement of the public (U.S. Environmental Protection Agency 1997, Bhattacharjee 2005, Cohn 2008). Important examples come from ornithological studies (Greenwood 2007, Kovács et al. 2008). Birds are good indicators of biodiversity generally, and they are easy to monitor because they are easy to identify and observe, and because there are many potential observers (National Audubon Society 2006, Greenwood 2007). Over the past decade, Cornell University has harnessed the enthusiasm of nonspecialist volunteers to explore questions such as the dynamics of infectious disease in bird populations and the impact of acid rain on their reproductive success. Those efforts have resulted in a list of peer-reviewed publications, clearly demonstrating the value of citizen science as a research tool (Hames et al. 2002, Altizer et al. 2004, Cohn 2008). Several other examples of published research confirm that nonspecialist volunteers can collect valid data (see, for instance, Evans et al. 2000, Fore et al. 2001, Lambert et al. 2005, Oberhauser et al. 2007, Delaney et al. 2008).

Volunteer participation in underwater monitoring presents unique challenges. Both terrestrial and marine projects require volunteer training but marine projects have the additional requirement of SCUBA diving skills. The last 20 years have seen a rapid increase in the numbers of recreational divers (Garrod and Gössling 2008), and research programs have begun to solicit divers as volunteers, making use of their natural interest in marine life. Among the research projects that developed the use of nonspecialist volunteers in marine monitoring, Coral Cay Conservation in Belize (Mumby et al. 1995), Fish Survey Project, conducted in Florida and the Caribbean (Pattengill-Semmens and Semmens 2003), and Reef Check, on a global scale (Hodgson 1999) are three significant examples. Coral Cay Conservation volunteers undergo an intensive eight-day training program in marine life identification and survey techniques. The training program incorporates lectures, practical exercises, individual tutoring, video, slides, and frequent testing. The course syllabus includes the identification of key species of macroalgae, seagrass, coral, and other marine invertebrates, as well as topographical features, species interaction, taxonomy, physiology, and consideration of coastal zone management issues and practices. After the training, volunteer divers conduct detailed survey transects for assessing marine resources for management initiatives. The Fish Survey Project assesses volunteers on fish species identification skills and classifies recruits as "beginners" or "experts" according to test results. Reef Check enrolls volunteers who pass a training course involving surveying techniques and diving skills. Participants perform successive surveys (fish, invertebrates, and substratum) at specific reef sites, transects and depths, following a strict protocol, and collect biophysical and socioeconomic data on that site under the guidance of professional scientists. Collectively these projects are able to involve few hundreds of recreational divers every year.

Asking volunteers to travel at their own expense to specific sites to perform surveys according to overbearing regimentation of the survey methods and strict protocols, may ensure uniform data collection, but carries the risk of making participation in the research project less attractive and so reducing the number of volunteers willing to participate. For detailed surveys, the use of volunteers would even be unsuitable. Detailed surveys require greater expertise in, for example, taxa identification, and an ability to maintain interest and accuracy. If demands are too great, people will not take part: the British Trust for Ornithology's Nest Sanitation project recruited very few participants, and thus reached no conclusions, because it required people to conduct such intensive work (Greenwood 2007). Darwall and Dulvy (1996) argue that the survey of "unknown" areas is sufficiently exciting for volunteers to maintain a high level of interest, but detailed studies repeated at a site lead to a significant drop in the level of interest, which is likely to lead to a loss in the quality of data collected. Striking the balance between work that is challenging enough to be satisfying but not so demanding as to off-put potential participants is not easy, especially because this balance varies for different people (Greenwood 2007). In an ideal world, all surveys would be conducted by a small team of highly experienced individuals but this is seldom possible due to lack of finance and time. Time is particularly important given the restricted physical limitations of diving surveys. For example, subtidal baseline surveys over large geographical scales require thousands of dives by hundreds of individuals, and this is most easily facilitated through the participation of a large number of volunteers (Darwall and Dulvy 1996).

There are also major educational and social benefits from the involvement of citizen volunteers in scientific projects. Participation in citizen-science projects provides a forum in which participants engage in thought processes similar to those that are part of science investigations, and increase their knowledge of ecology and environmental issues (Trumbull et al. 2000, Evans and Birchenough 2001, Brossard et al. 2005). The "self-education" of those collecting data, "the raising of a conservation force for change," and the pride that citizen scientists take in helping advance scientific knowledge and protecting the environment are also recognized benefits (Cohn 2008).

Since 1999, in an effort to maximize recreational diver participation, we have been testing a method of volunteer involvement that ensures reliability but does not diminish the diver enjoyment (i.e., without changing the normal recreational dive profile: depth, time, path; Goffredo et al. 2004). We wanted to give people an opportunity to become involved in environmental conservation in a novel way, balancing the need to collect good quality data with public education. This effort has therefore been to unite research with recreation, putting citizens at the forefront of the conservation drive. We first designed the "Mediterranean *Hippocampus* Mission," that focused on only one taxon: seahorses (Goffredo et al. 2004). Approximately 2500 recreational divers took part in the search for seahorses, and reported sightings via a user-friendly questionnaire. Volunteers enabled us to map the distribution of seahorses in the Italian Mediterranean Sea. This achievement prompted us to design a more ambitious project, named "Divers for the Environment: Mediterranean Underwater Biodiversity Project," the subject of this paper. The aims of Divers for the Environment were:

1) Involving as many people as possible in biodiversity monitoring;

2) Validating this new volunteer based monitoring approach, where volunteers perform recreational dives (i.e., pre-oriented precise transects are not carried out), and comparing results with those from professional investigations;

3) Developing a volunteer sightings-based index model for evaluating the status of the marine environment;

4) Making information available to the whole community by wide dissemination of the results.

The dissemination of information from citizen science projects can go far beyond the participants themselves. The mass media are keen to report findings of studies involving citizen-volunteers (Evans et al. 2000, Foster-Smith and Evans 2003, Goffredo et al. 2004). Evans et al. (2000) suggested that, because of media attention, the results of volunteer surveys may have wider impacts than other "purely scientific" studies. Wider implications are far-reaching because there can be little doubt that the public's failure to comprehend scientific issues is a root cause of the under-funding of science (Foster-Smith and Evans 2003). Citizen volunteers may also bring attributes of scientific studies, such as special skills (Foster-Smith 2000), specialist knowledge (Harrison et al. 1998) and new insights (Kendall and Lewis 1986), so that they contribute significantly more than a workforce that collects data (Foster-Smith and Evans 2003).

Materials and Methods

Survey questionnaires

From 2002 to 2005, we asked recreational divers to complete a questionnaire recording the presence of animal and plant taxa and refuse (litter). The questionnaire had two sections: one with photographs to identify the surveyed taxa (Appendix A: Fig. A1), the other with a form to record data (Appendix B: Fig. B1).

Sixty-one organismal taxa were surveyed (four vegetal taxa and 57 animal taxa; Appendix B: Fig. B1). It was necessary to have a long taxa list to address the overarching aim of assessing the quality of the environment from its biodiversity status (i.e., a single species by itself was not considered as an environmental quality indicator; Grime 1997, Therriault and Kolasa 2000). In a census of a comparable number of taxa (56 reef taxa), Darwall and Dulvy (1996) show that nonspecialist volunteer divers were able to reach a level of precision equivalent to an experienced researcher. Surveyed taxa had to be previously well known by volunteer recreational divers or easily recognizable (see Appendix C for volunteer training methods), benthic (highly mobile pelagic species were not censused; after Darwall and Dulvy 1996), historically expected to be found throughout the entire Mediterranean Sea (based on Riedl 1991 and the databases Global Biodiversity Information Facility, Ocean Biogeographic Information System, and MarineSpecies) and representative of each of the major trophic levels (databases available online).[4,5,6] These characteristics were necessary in order that the method is suitable for amateurs and tasks are realistic and achievable (Oliver and Beattie 1993, Pearson 1994, Therriault and Kolasa 2000, Foster-Smith and Evans 2003, Greenwood 2003, Newman et al. 2003, Goffredo et al. 2004, Bell 2007, Cohn 2008), the variation in biodiversity composition detected among geographic areas is not solely attributable to natural variation (Pearson 1994), and the estimated level of biodiversity is related to local conditions. The relevance of each taxon in revealing variation in diversity among sites was quantified using the "global BEST test" (Bio-Env + STepwise; PRIMER-E version 6 software, PRIMER-E, Ltd., Ivybridge, UK; Clarke et al. 2008), in order to determine the minimum subset of taxa which would generate the same multivariate sample pattern as the full assemblage.

As in previous works (Schmitt and Sullivan 1996, Pattengill-Semmens and Semmens 2003, Goffredo et al. 2004), the required data were general information about the surveyor, level of diving qualification, diving agency that issued the license, technical information about the dive (place, date, time of day, depth, length of time), type of habitat explored (rocky bottom, sandy bottom, or other habitat), and an estimate of the abundance of surveyed organisms (Appendix B: Fig. B1). For each taxon we defined the scale of abundance as "rare," "frequent," or "abundant" based on the frequency at which the taxon is normally encountered. This frequency was estimated using scientific databases, literature, and personal observations. As an example, 1–4 rainbow wrasse was classed as rare, 5–10 as frequent, and more than 10 as abundant. Litter (fish pots, nets, or general refuse) was also recorded.

The diving certification level of volunteers ranged from open water divers (at least six recorded dives), to instructors (at least 100 recorded dives). The diving certification level was ranked on an ordinal scale, based on the international standards (World Recreational Scuba Training Council [WRSTC] or World Confederation of Underwater Activities [CMAS]): open water diver (level 1), advanced diver (level 2), rescue diver (level 3), divemaster (level 4), instructor (level 5).

Simple random sampling design was used (i.e., volunteer divers were not forced; they performed survey dives when and where it was convenient for them). Also the recreational dive profile (dive depth, time, path, and safe diving practices) was not modified for the surveys: divers performed the dive as they normally do during sport diving (after Goffredo et al. 2004). This was because the aim of the study was to test the validity of using data from recreational dives for marine monitoring. During the survey dive each diver was responsible for observing plants, invertebrates and fishes, as well as litter. Soon after the dive, each participant completed a recording questionnaire (i.e., number of recorded questionnaires = number of dives performed). The completion of data questionnaire shortly after the dive, and the assistance of trained professional divers during data recording were key elements of the survey protocol to control data quality (Goffredo et al. 2004). Divemasters and other trainers that worked with the volunteers all attended the training courses for professional divers (see Appendix C). Their similar backgrounds and training assured limited influence on the accuracy of the volunteers under their supervision.

Assessing characteristics of sites: the survey station parameters

Incomplete or illegible questionnaires were discarded, as were those that demonstrated misunderstanding of methods (for example, multiple dives recorded on the same questionnaire), amounting to 16.6% of questionnaires submitted.

Data were aggregated according to type of habitat explored: rocky bottom, sandy bottom or other. We calculated the marine biodiversity index (V.MBI) for rocky bottom sites, since this environment was recorded in the highest number of survey questionnaires, enabling spatiotemporal comparison of results. Data from sites that did not have rocky bottoms were not used for any of the analyses in this paper. The questionnaires from rocky habitats were aggregated by dive site. We used the term "survey station" to define a dive site that produced at least 10 valid questionnaires in one year. Questionnaires from the survey stations were defined as "useful questionnaires" and were statistically analyzed. Dive sites that failed to reach the quorum of ten valid questionnaires over one year were defined as "sparse sites" and their questionnaires, defined as "sparse questionnaires," were not elaborated.

As in previous studies (Schmitt and Sullivan 1996, Pattengill-Semmens and Semmens 2003, Goffredo et al. 2004), we performed a statistical analysis for each survey station by calculating the following parameters: number of useful questionnaires recorded in one year; mean date, time of day, and depth of survey; number of vegetal (S_V) and animal (S_A) sighted taxa (aggregated over all questionnaires); sighting frequency of each taxon (%SF; expressed as percentage of dives in which the taxon was sighted); relative abundance of each taxon (abundance score, calculation follows); biodiversity values, vegetal (V) and animal (A) biodiversity, calculated by the Shannon-Wiener index (observed biodiversity H_{SH}, maximum biodiversity $L(S)$, equipartition index E_{SH}; Magurran 1988) using the relative abundance of each taxon (abundance score) to calculate the parameter p_i of the Shannon-Wiener index ($p_i =$

---

[4] ⟨http://www.gbif.org/⟩

[5] ⟨http://iobis.marine.rutgers.edu/⟩

[6] ⟨http://www.marinespecies.org/⟩

proportion of individuals of the taxon *i*; Magurran 1988); litter sighting frequency (%LF) expressed as percentage of dives where litter was observed.

To calculate the abundance score, we first calculated density score $= [(R \times 1) + (F \times 2) + (A \times 3)]/n$ where R, F, and A are the number of times the taxon was recorded as "rare," "frequent," or "abundant," respectively; 1, 2, and 3 are normalized abundance values assigned to the classes "rare," "frequent," and "abundant"; and $n = (R + F + A)$ (for statistical characteristics and rationale please see Schmitt and Sullivan 1996, Pattengill-Semmens and Semmens 2003). Then abundance score = density score × %SF (for statistical characteristics and rationale please see Schmitt and Sullivan 1996, Pattengill-Semmens and Semmens 2003).

Construction of the biodiversity evaluation model

Preliminary remarks.—In our model, the measure of biodiversity at a single survey station derives from the overall recorded information on censused taxa; single taxa by themselves are not considered indicators of general patterns (Grime 1997, Therriault and Kolasa 2000). The observed marine biodiversity has been synthesized into components of the Shannon-Wiener index (Magurran 1988, Lohrer et al. 2004).

To evaluate the biodiversity level at each survey station, we made a comparison between the values of parameters for each station and those calculated for a virtual "reference station." The parameters were S_{V}, H_{SHV}, E_{SHV}, S_{A}, H_{SHA}, E_{SHA} and %LF, defined as "main parameters," and sighting frequencies of individual taxa, defined as "special parameters." The virtual reference station was only one for the entire study. The assumption was that the virtual reference station represented the best current condition for a station in a rocky bottom habitat (i.e., its parameters were calculated from the actual stations having the best parameter conditions: higher biodiversity, lowest presence of litter). The parameter values of each individual station were expected to match those of the virtual reference station; otherwise they were considered as "penalties." The number of penalties resulting in the individual station determined the biodiversity index value.

Parameter calculation of the virtual reference station.—We calculated the virtual reference station parameter values as follows:

1) We calculated the "main" and "special" parameters of each survey station from the total number of useful questionnaires obtained during the four years.

2) For each of the parameters (main and special) we calculated the mean value among the stations and lower 95% confidence limit (upper 95% confidence limit for %LF).

3) We compared the parameter values of each station with the confidence limits obtained. If a value was below (above, for %LF), this counted as a "non-matching point" for the station. We summed the number of non-matching points for the station.

4) We calculated the mean number of non-matching points per station and the 95% upper confidence limit. We rejected the stations with more non-matching points than the confidence limit.

5) For the stations remaining after the rejection we returned to step 2. The 2, 3, and 4 cycle was repeated until all the remaining stations had a number of non-matching points less than or equal to the upper confidence limit.

6) We assumed as the critical values for the virtual reference station the lower 95% confidence limits of the means for the remaining stations (upper 95% limit for %LF).

Index (V.MBI [volunteers marine biodiversity index]).—For each year, we compared the values of the parameters for each station with the values of the virtual reference station. The parameters that did not reach the minimum requirements were considered as penalties (for S_{V}, H_{SHV}, E_{SHV}, S_{A}, H_{SHA}, and E_{SHA} and the special parameters, the value had to be equal or higher than that of the virtual reference station; for the %LF, the value had to be equal or lower than that of the virtual reference station). Each penalty was assigned a value calculated according to the frequency with which the penalty itself occurred in the totality of the stations: penalty value = 100 − penalty frequency (i.e., the percentage of stations in which the penalty was present). The sum of the penalty values was calculated for the main parameters and for the special parameters (we got two sums). Each sum was normalized on a scale from 0 to −1, where 0 indicated the absence of penalties and −1 indicated all penalties. We calculated marine biodiversity index for each individual station as the mean of the two normalized sums. The index was reduced to five classes: very good (for values between 0 and −0.125), good (−0.126 to −0.375), mediocre (−0.376 to −0.625), low (−0.626 to −0.875), and very low (−0.876 to −1).

Assessment of the validity of data collected by nonspecialist volunteers

Validation trials.—Comparisons were made between records from trained volunteers and independent records from a marine biologist (over 2000 hours of marine surveying experience), hereafter referred to as the "control diver." The explanations for the experimental design comparing volunteers to the control diver are after Mumby et al. (1995) and Darwall and Dulvy (1996):

1) The control diver was the same individual for all validations; in each validation the volunteer divers were different from previous ones (i.e., each volunteer was tested only once);

2) The control diver dived simultaneously with trained volunteers without interfering with them;

3) Validation dive sites were not selected prior to the assessment; the control diver dived where the diving

Table 1. Definition and derivation of terms used to describe components of the accuracy and consistency of volunteers data.

Parameter	Definition and derivation of parameter
Accuracy	Similarity of volunteer-generated data to reference values from a control diver measured as rank correlation coefficient and expressed as a percentage in the text. This measure of accuracy is assumed to encompass all component sources of error.
Consistency	Similarity of data collected by separate volunteers during the same dive. This was measured as rank correlation coefficient and expressed as percentage in the text. This measure of consistency is assumed to encompass all component source of error.
Percent identified	The percentage of the total number of taxa present that were recorded by the volunteer diver. The total number of taxa present was derived from the control diver data (i.e., we assumed the taxa recorded by the control diver to be all the taxa present).
Correct identification	The percentage of volunteers that correctly identified individual taxa when the taxon was present.
Correctness of abundance ratings (CAR)	This analysis quantified the correctness in abundance ratings made by the volunteer. It has been expressed as the percentage of the 62 surveyed taxa whose abundance has been correctly rated by the volunteer (i.e., the value of the rating indicated by the volunteer was equal to the reference value recorded by the control diver).

Note: Modified from Mumby et al. (1995).

center officer planned the dive for that day, accordingly to safe conditions (weather, currents, divers experience);

4) All trials were conducted in between 09:00 and 16:00 to avoid changes in activity between nocturnal and diurnal taxa populations;

5) At the end of the dive the control diver filled the questionnaire independently and apart from the volunteers without any interference with volunteer data recording;

6) For each trial an inventory of taxa (with abundance rating) was generated by the control diver, and this was compared with the inventory generated by each volunteer surveyor to identify data accuracy.

Data validation statistics.—Correlation analyses between the records of the control diver and the records of the volunteers were performed to assess agreement between the independent records (Darwall and Dulvy 1996, Evans et al. 2000). This comparison was performed each year at different survey stations with different volunteers, to constantly monitor the validity of the data collected and the effectiveness and consistency of the annual training workshops. A variety of nonparametric statistical tests were used to analyze the survey data:

1) Spearman rank correlation coefficients (ρ_s) were calculated and results displayed in terms of mean value and 95% confidence limit. Several terms were used to describe sources of inaccuracy, error and variation in survey data (Table 1).

2) Cronbach's alpha (α) correlation was used to analyze the reliability of survey data (Hughey et al. 2004). The α coefficient is a calculated value (ranging between 0 and 1, and expressed as a percentage in the text) based on the average correlation of items within a test if the response categories are standardized (Coakes and Steed 1997). Values above 0.5 are considered acceptable as evidence of a relationship (Nunnally 1967, Hair et al. 1995), an α above 0.6 is considered an effective reliability level (Flynn et al. 1994), while values above 0.7 are more definitive (Peterson 1994). The α coefficient was calculated for each volunteer taxa inventory against the control diver inventory. The results were displayed in terms of mean value and 95% confidence limit.

3) Czekanowki's proportional similarity index SI was used to obtain a measure of similarity between each volunteer and the control diver ratings (as for Sale and Douglas [1981] and Darwall and Dulvy [1996]):

$$\mathrm{SI}_{ij} = 1 - \frac{1}{2}\sum_{n=1}^{s}[p_{in} - p_{jn}]$$

where there are s taxa, and p_{in} and p_{jn} represent the proportions of individuals in census i and j respectively that belong to the nth species. The value $p_{in} - p_{jn}$ is taken as the absolute difference between the two proportions. The index ranges from 0 when two censuses have no taxa in common to 1 when the distribution of abundance ratings across species is identical. Values above 0.5 are considered as indication of sufficient levels of precision, while values above 0.75 are considered as high levels of precision (Darwall and Dulvy 1996). The results were displayed in terms of mean value and 95% confidence limit.

To develop eligibility criteria for future surveys, we identified independent variables (diving certification level and group size of participants) to examine their effect on the precision of volunteers. The possible influence of dive time and depth on volunteer precision was also assessed. For all of these analyses the Spearman rank correlation was tested.

Statistical analyses were conducted using SPSS 12.0 for Windows (SPSS, Chicago, Illinois, USA).

Results

Quality of recreational volunteer-generated data: the validation trials

The overall trends of accuracy, consistency, reliability, and similarity are described, including an inspection of the individual components of accuracy (defined in Table 1) and species-level analysis.

Table 2. Quality of volunteer-generated data; results of the 38 validation trials performed during the four-year research project (2002–2005).

Station name	Code	Date	Team size†	Cert. level‡	Depth (m)	Dive time (minutes)
2002						
Gorgonie	gr-14	25 Apr	9	3.0 (2.1–3.9)	21 (19–22)	42 (41–43)
Punta della Madonna	pm-16	2 Jun	7	2.4 (1.6–3.3)	26 (20–32)	37 (32–42)
Scogliera Parco Marino	spm-31	15 Jun	7	2.3 (1.3–3.3)	4 (4–5)	63 (58–69)
Tato Point	tp-14	22 Jun	10	1.7 (1.3–2.1)	28 (26–30)	43 (40–47)
Calafuria	c-14	23 Jun	10	1.8 (1.0–2.6)	13 (11–16)	58 (54–62)
Ancorone	a-14	24 Aug	6	1.5 (0.8–2.2)	17 (15–19)	46 (43–49)
Gorgonie	gr-14	25 Aug	9	1.4 (0.9–2.0)	17 (15–18)	40 (40–41)
Tato Point	tp-14	25 Aug	10	1.4 (1.0–1.8)	18 (16–19)	43 (42–44)
Scoglione	s-15	4 Oct	4	2.7 (1.6–3.8)	16 (14–17)	49 (42–56)
Secca Turco	st-15	4 Oct	5	3.0 (2.4–3.6)	23 (20–25)	44 (40–48)
Scoglione	s-15	5 Oct	7	1.6 (0.8–2.3)	14 (13–15)	56 (52–59)
Secca Turco	st-15	5 Oct	7	2.7 (2.2–3.3)	25 (22–27)	37 (35–39)
2003						
Cartellino	ct-14	11 May	4	2.3 (1.3–3.2)	22 (21–22)	49 (46–51)
Calafuria	c-14	18 May	6	2.0 (1.1–2.9)	10 (7–13)	45 (44–46)
Cala Fetente	cf-24	23 May	6	2.3 (1.5–3.2)	8 (6–9)	33 (30–36)
C.po Spartivento	cs-24	24 May	6	3.0 (2.0–4.0)	22 (16–27)	43 (41–44)
Grotta Azzurra	ga-24	24 May	11	2.5 (1.6–3.3)	16 (13–19)	47 (43–52)
Civitata	cv-15	7 Jun	7	1.4 (0.8–2.0)	11 (11–12)	50 (50–51)
Formiche	f-15	8 Jun	5	1.4 (0.6–2.2)	13 (12–15)	50 (46–54)
Forbici	fr-16	4 Jul	15	2.1 (1.4–2.7)	17 (15–19)	49 (44–53)
Picchi Pablo	pp-16	5 Jul	9	2.7 (1.9–3.4)	18 (15–22)	44 (35–52)
Sc. Remaiolo	sr-16	26 Jul	6	1.0	17 (15–18)	42 (40–43)
Secca di Fonza	sdf-16	26 Jul	6	1.0	17 (16–19)	39 (39–40)
Spiaggia di Portoazzurro	spa-16	7 Nov	11	1.5 (0.8–2.1)	7 (6–8)	30 (29–31)
2004						
P.ta della Fica	pf-15	28 May	6	2.3 (1.7–3.0)	16 (12–20)	42 (41–42)
Formiche	f-15	30 May	10	1.5 (0.9–2.1)	13 (12–14)	47 (45–49)
Calafuria	c-14	13 Jun	14	1.5 (0.9–2.0)	7 (6–8)	38 (38–39)
Sc. Remaiolo	sr-16	23 Jul	12	1.8 (1.0–2.5)	12 (11–13)	44 (42–47)
Corbelli	cri-16	24 Jul	19	1.5 (1.0–2.0)	12 (11–13)	47 (45–48)
Sc. Remaiolo	sr-16	24 Jul	18	1.5 (1.0–2.0)	12 (11–12)	51 (50–52)
C.po Focardo	cf-16	27 Jul	10	1.6 (0.8–2.4)	7 (6–8)	43 (42–43)
Cannelle	cn-16	27 Nov	8	1.8 (0.8–2.7)	10 (7–13)	40 (37–43)
Picchi Pablo	pp-16	28 Nov	13	1.5 (0.9–2.1)	10 (9–11)	47 (42–53)
2005						
Cala Turchi	ct-30	27 Oct	3	4.2 (3.3–5.0)	23 (20–27)	46 (43–48)
Punta Secca di Caprara	psc-30	27 Oct	3	3.5 (2.0–5.0)	27 (20–33)	46 (43–50)
Spiaggia di Portoazzurro	spa-16	29 Oct	9	1.7 (0.8–2.5)	8 (7–9)	45 (43–47)
Sc. Remaiolo	sr-16	30 Oct	10	1.6 (0.8–2.4)	13 (11–15)	46 (39–52)
Cala Caffè	cc-30	31 Oct	5	3.5 (2.3–4.7)	21 (18–23)	45 (45–46)

Notes: Parameter definitions are in Table 1 and in *Materials and methods*. Values in parentheses are 95% CIs. Volunteers tested in 2002 had <1 year of survey experience, those tested in 2003 had <2 years of survey experience, those tested in 2004 had <3 years of survey experience, and those tested in 2005 had <4 years of survey experience.

† Number of volunteers.

‡ Diving certification level of volunteers.

Thirty-eight validation trials were performed (Table 2). A total of 324 different volunteers were tested, with a mean number of volunteers per validation team of 9 (95% CI = 7–10). Mean diving certification level of volunteers varied significantly among teams from 1.0 to 4.2 (Table 2).

There was significant variability in the accuracy of validation trials. The mean accuracy of each team ranged from 38% to greater than 90%, with the majority of teams (76%) performed with mean accuracy of between 50% and 80% (Table 2). Intra-group variation was approximately 21% (coefficient of variation, CV) per team. Accuracy was not correlated with volunteers diving certification level ($\rho_s = -0.262$, $N = 38$, $P = 0.112$), number of participants in the trial group ($\rho_s = -0.110$, $N = 38$, $P = 0.511$), depth of the trial ($\rho_s = -0.281$, $N = 38$, $P = 0.087$), or dive time of the trial ($\rho_s = -0.025$, $N = 38$, $P = 0.882$). A consistent trend emerged from the regression analysis between time from the beginning of the trials and accuracy, which indicated an increase in accuracy of 7 points each year ($\rho_s = 0.702$, $N = 38$, $P < 0.001$; Accuracy (%) = 7.013time (in years) + 57.465).

Consistency showed a similar pattern to that of accuracy; the mean consistency of each team ranged from 39% to 91%, with the majority of teams (76%) performing with a mean consistency of between 50% and 80% (Table 2). Intra-group variation was at approximately 26% (CV) per team. Consistency was not

TABLE 2. Extended.

Accuracy	Consistency	Percent identified	CAR	Reliability (α)	Similarity index
62.5 (53.3–71.7)	43.4 (38.5–48.4)	67.5 (60.5–74.5)	81.7 (78.4–85.0)	75.7 (66.6–84.8)	59.7 (52.2–67.1)
42.7 (34.6–50.8)	44.3 (36.3–52.2)	64.8 (47.8–81.9)	72.8 (69.3–76.4)	55.1 (47.2–63.0)	44.1 (37.2–51.0)
57.6 (50.0–65.2)	52.3 (47.8–56.7)	63.8 (49.0–78.6)	80.6 (78.7–82.6)	68.8 (58.1–79.5)	55.1 (43.4–66.7)
54.2 (48.7–59.6)	61.9 (58.3–65.4)	58.5 (53.3–63.6)	79.5 (77.7–81.3)	77.3 (73.5–81.1)	57.8 (54.4–61.2)
54.8 (50.6–58.9)	49.5 (44.2–54.8)	65.3 (58.6–72.0)	76.0 (73.6–78.3)	64.0 (55.7–72.3)	52.4 (46.6–58.3)
70.4 (54.2–86.5)	65.4 (56.3–74.5)	79.5 (72.0–86.9)	84.1 (76.3–92.0)	78.2 (62.8–93.7)	67.4 (49.6–85.1)
69.8 (58.1–81.4)	58.2 (51.8–64.6)	83.3 (76.3–90.4)	85.3 (78.9–91.7)	82.7 (75.0–90.4)	65.7 (53.0–78.4)
66.1 (56.8–75.5)	60.5 (56.0–65.0)	78.0 (68.0–88.0)	82.4 (76.4–88.5)	81.6 (76.3–87.0)	63.0 (54.8–71.1)
57.6 (40.7–74.4)	48.5 (43.7–53.3)	75.0 (58.7–91.3)	82.3 (70.0–94.5)	77.4 (62.6–92.2)	51.3 (28.9–73.8)
49.0 (39.8–58.1)	49.3 (42.4–56.2)	60.0 (46.1–73.9)	80.6 (78.9–82.4)	69.9 (60.0–79.7)	50.4 (40.3–60.6)
38.4 (26.4–50.4)	39.0 (28.5–49.5)	57.1 (39.9–74.4)	73.3 (68.9–77.6)	52.2 (35.3–69.1)	39.0 (29.5–48.4)
53.8 (47.0–60.6)	50.6 (43.9–57.4)	54.0 (45.2–62.8)	85.7 (83.2–88.2)	77.4 (67.2–87.5)	56.3 (46.7–66.0)
68.5 (53.0–84.0)	60.8 (50.0–71.5)	77.3 (58.0–96.5)	67.7 (59.1–76.4)	79.7 (66.7–92.8)	67.6 (54.7–80.6)
80.7 (63.6–97.9)	56.1 (45.1–67.1)	85.2 (71.8–98.6)	89.0 (80.3–97.7)	79.5 (64.0–95.0)	66.8 (46.3–87.2)
68.0 (57.4–78.6)	49.5 (41.3–57.7)	70.8 (55.8–85.9)	94.1 (92.1–96.0)	84.5 (73.2–95.8)	63.1 (50.7–75.5)
67.0 (55.2–78.8)	61.1 (56.5–65.7)	72.0 (60.4–83.6)	74.7 (68.2–81.2)	82.9 (76.1–89.7)	70.5 (60.9–80.1)
52.3 (44.9–59.7)	57.0 (53.4–60.6)	73.9 (67.9–79.8)	68.3 (63.9–72.8)	66.9 (60.6–73.1)	54.1 (48.9–59.3)
90.1 (87.2–93.1)	90.5 (88.5–92.5)	93.2 (91.3–95.1)	92.6 (88.9–96.4)	94.7 (92.3–97.0)	88.9 (84.3–93.4)
67.7 (65.2–70.2)	74.9 (69.7–80.2)	77.9 (72.8–82.9)	73.5 (70.3–76.8)	79.5 (77.3–81.6)	66.5 (63.6–69.5)
61.5 (55.8–67.1)	55.0 (52.7–57.4)	67.4 (60.1–74.6)	73.1 (70.4–75.8)	72.7 (67.2–78.1)	58.6 (53.9–63.3)
59.0 (52.3–65.6)	51.5 (46.1–56.8)	71.4 (61.3–81.6)	73.8 (70.0–77.7)	73.0 (66.7–79.3)	56.7 (50.4–62.9)
80.1 (70.1–90.1)	76.4 (70.0–82.8)	86.1 (78.3–93.9)	84.1 (76.4–91.9)	86.7 (78.7–94.7)	76.8 (66.9–86.8)
74.3 (54.6–94.1)	57.9 (47.9–68.0)	76.4 (55.8–97.0)	84.7 (73.8–95.6)	83.3 (68.4–98.3)	74.0 (53.8–94.2)
72.7 (59.3–86.0)	54.2 (47.6–60.8)	64.8 (47.7–81.9)	90.8 (86.9–94.7)	80.6 (68.6–92.6)	65.2 (49.2–81.2)
68.1 (59.7–76.4)	62.8 (56.9–68.7)	64.6 (56.4–72.7)	81.7 (77.3–86.2)	83.2 (75.9–90.4)	65.5 (57.7–73.3)
69.4 (64.8–74.0)	65.8 (61.1–70.4)	75.6 (68.3–82.9)	73.9 (72.3–75.5)	81.5 (78.4–84.7)	66.5 (62.5–70.5)
63.1 (55.8–70.5)	72.0 (69.0–74.9)	62.2 (55.6–68.9)	84.2 (81.6–86.8)	82.6 (77.5–87.6)	64.9 (57.9–71.8)
68.6 (62.3–74.9)	63.3 (59.8–66.8)	80.8 (73.0–88.5)	77.0 (70.7–83.3)	81.5 (76.7–86.4)	64.7 (57.2–72.3)
71.2 (63.3–79.1)	61.3 (58.9–63.7)	74.6 (68.3–80.8)	80.6 (75.4–85.9)	83.1 (77.9–88.4)	70.0 (62.6–77.4)
76.0 (70.3–81.8)	65.9 (63.7–68.1)	85.8 (81.2–90.3)	80.8 (76.7–85.0)	85.7 (81.3–90.1)	73.7 (67.9–79.4)
84.7 (78.9–90.6)	81.2 (77.9–84.6)	85.2 (80.5–89.9)	87.3 (82.2–92.3)	90.9 (87.2–94.6)	81.5 (75.6–87.5)
78.6 (62.7–94.4)	64.6 (56.0–73.2)	84.2 (74.3–94.0)	86.7 (78.2–95.2)	84.4 (69.7–99.2)	77.7 (61.8–93.5)
73.4 (61.6–85.2)	64.4 (60.2–68.7)	74.8 (60.8–88.9)	75.7 (68.0–83.3)	82.6 (74.7–90.5)	68.3 (56.1–80.5)
80.6 (63.6–97.6)	67.5 (55.4–79.7)	79.6 (59.3–100.0)	85.5 (77.5–93.4)	92.6 (87.1–98.2)	80.8 (68.4–93.1)
88.5 (77.9–99.1)	74.6 (66.2–82.9)	84.1 (68.3–100.0)	88.2 (82.6–93.7)	94.9 (89.9–100.0)	85.0 (73.6–96.4)
75.3 (66.0–84.6)	71.4 (66.6–76.1)	76.3 (69.4–83.2)	87.1 (83.0–91.1)	85.2 (76.5–93.9)	73.2 (65.3–81.1)
74.4 (64.0–84.8)	71.7 (67.7–75.6)	77.9 (69.6–86.1)	94.6 (90.8–98.4)	83.8 (76.3–91.3)	71.5 (61.3–81.6)
82.0 (69.8–94.2)	68.3 (60.3–76.4)	85.7 (73.5–97.9)	86.5 (77.7–95.2)	91.1 (83.2–99.0)	83.3 (71.7–94.8)

correlated with depth of the trial ($\rho_s = -0.209$, $N = 38$, $P = 0.209$), the dive time of the trial ($\rho_s = 0.094$, $N = 38$, $P = 0.574$), or number of participants in the group ($\rho_s = 0.021$, $N = 38$, $P = 0.899$). Interestingly, there was an inverse correlation between volunteers diving certification level and consistency ($\rho_s = -0.372$, $N = 38$, $P = 0.022$). The regression analysis between time from the beginning of the trials and consistency showed a consistent trend with an increase of 6 points in consistency each year ($\rho_s = 0.680$, $N = 38$, $P < 0.001$; consistency (%) = 5.798[time (in years)] + 52.657).

Most survey teams managed to correctly identify approximately 75% of the taxa present in each survey trial (87% of the teams correctly identified a mean percentage of between 60% and 90%; Table 2). Intra-group variation was approximately 20% (CV) per team. The ability to correctly identify taxa was not correlated with the diving certification level of the team members ($\rho_s = -0.275$, $N = 38$, $P = 0.095$), the group size of participants ($\rho_s = -0.157$, $N = 38$, $P = 0.348$), depth ($\rho_s = -0.132$, $N = 38$, $P = 0.430$) or dive time of the trial ($\rho_s = 0.143$, $N = 38$, $P = 0.392$).

A positive correlation between the number of validation trials in which the taxon was present and the level of correct identification by volunteers was detected (Table 3; $\rho_s = 0.448$, $N = 46$, $P < 0.01$; correct identification (%) = 1.057[presence frequency] + 53.952). Sixteen rare taxa were not present (i.e., were not recorded by the control diver) in any of the 38 validation trials, thus assessment of correct identification was not possible.

Most survey teams managed to correctly rate the abundance of approximately 82% of the surveyed taxa (95% of the teams produced a mean correctness of abundance ratings, CAR, of between 70% and 100%; Table 2). Intra-group variation was approximately 9%

TABLE 3. Taxon-level analyses.

Taxon		Correct identification (%)			
Common name	Scientific name	Mean	95% CI	*N*	Best taxon
Mermaid's wine glass	*Acetabularia acetabulum*	94.5	89.7, 99.4	12	
Damselfish	*Chromis chromis*	94.1	90.7, 97.6	35	×
Mediterranean tapeweed	*Posidonia oceanica*	93.6	87.2, 100.0	19	×
Sea anemone	*Anemonia viridis*	91.8	86.0, 97.6	10	
Salema	*Sarpa salpa*	91.0	85.1, 96.9	20	
Yellow cluster anemone	*Parazoanthus axinellae*	89.2	84.1, 94.3	18	
Precious red coral	*Corallium rubrum*	87.5	75.0, 100.0	6	
Red gorgonian	*Paramuricea clavata*	87.0	74.1, 100.0	3	×
Other fishes		86.5	79.5, 93.4	27	
Dusky grouper	*Epinephelus guaza*	84.0	74.6, 93.3	7	×
Fin shell	*Pinna nobilis*	83.3	66.7, 100.0	4	
Other bivalves		83.3	66.7, 100.0	2	×
Moray eel	*Muraena helena*	83.3	71.5, 95.0	9	
Other sponges		82.2	75.2, 89.2	31	
Fan tube worm	*Sabella spallanzanii*	81.4	71.4, 91.4	17	
Other sea stars		78.4	66.9, 89.9	16	×
Dotted sea slug	*Peltodoris atromaculata*	78.2	56.4, 100.0	8	
Petrosia	*Petrosia ficiformis*	77.5	67.7, 87.4	15	
Other echinoids		77.2	69.2, 85.4	27	
Sea lace	*Sertella septentrionalis*	76.3	66.7, 86.0	15	×
Sea rose	*Peyssonnelia squamaria*	76.3	67.5, 85.1	25	
Common spiny lobster	*Palinurus elephas*	75.8	51.6, 100.0	3	
Rainbow wrasse	*Coris julis*	75.6	66.9, 84.2	31	
False coral	*Myriapora truncata*	75.2	65.5, 84.9	26	×
Other octocorals		74.2	62.8, 85.5	15	
Sea raven	*Sciaena umbra*	74.1	51.4, 96.9	2	
Sea red potato	*Halocinthya papillosa*	72.9	63.2, 82.7	23	×
Other vegetals		68.8	58.4, 79.3	28	×
Litter		67.5	56.5, 78.5	17	
Sea lily	*Antedon mediterranea*	66.7		1	
Other sedentary worms		65.4	51.9, 78.9	23	×
Brain sponge	*Chondrilla nucula*	59.3	31.1, 87.5	5	×
Common octopus	*Octopus vulgaris*	59.2	41.2, 77.3	2	
Other holoturians		59.1	47.0, 71.2	17	×
Cuttlefish	*Sepia officinalis*	57.1		1	
Spider crab	*Maja squinado*	57.1		1	
Other gastropods		56.0	37.5, 74.5	10	×
Other ascidians		53.5	31.0, 76.0	5	×
Other hexacorals		46.2	32.3, 60.1	21	
Other bryozoans		34.3	11.9, 56.6	9	×
Pencil sea urchin	*Stylocidaris affinis*	33.3		1	
Cerianthid anemone	*Cerianthus membranaceus*	32.6	9.2, 56.1	4	
Other decapods		21.4	−20.6, 63.4	2	
Sea cucumber	*Stichopus regalis*	14.3		1	
Other ophiuroids		14.3		1	
Pentagon sea star	*Ceramaster placenta*	0.0		1	
Eyed electric ray	*Torpedo torpedo*			0	
Smooth brittlestar	*Ophioderma longicaudum*			0	×
Thornback ray	*Raja clavata*			0	×
Anglerfish	*Lophius piscatorius*			0	
John dory	*Zeus faber*			0	
Flying gurnard	*Dactylopterus volitans*			0	×
Winged oyster	*Pteria hirundo*			0	×
Purple dye murex	*Bolinus brandaris*			0	×
Red dead man's fingers	*Alcyonium palmatum*			0	×
Box crab	*Calappa granulata*			0	×
Giant tun	*Tonna galea*			0	×
Long-snouted branched seahorse	*Hippocampus ramulosus*			0	×
Short-snouted seahorse	*Hippocampus hippocampus*			0	
European lobster	*Homarus gammarus*			0	
Other crinoids				0	×
Other cephalopods				0	×

Notes: Correct identifications were generated from a maximum sample size of 38 validation trials performed at the stations listed in Table 2, from 25 April 2002 to 31 October 2005. *N* is the actual sample size for each taxon (i.e., presence frequency, the number of validation trials in which the taxon was present). Refer to Table 1 for definition of "correct identification." Best taxon refers to a subset of 27 taxa. The BEST test (Bio-Env + STepwise; PRIMER-E version 6 software) was performed on the total sample size of 16 533 questionnaires collected over the four years of research. These 27 taxa constituted the minimum subset that generated the same multivariate sample pattern derived from the full taxa assemblage and represented in Fig. 4 (BEST test, $\rho_s = 0.951$, $P < 0.01$).

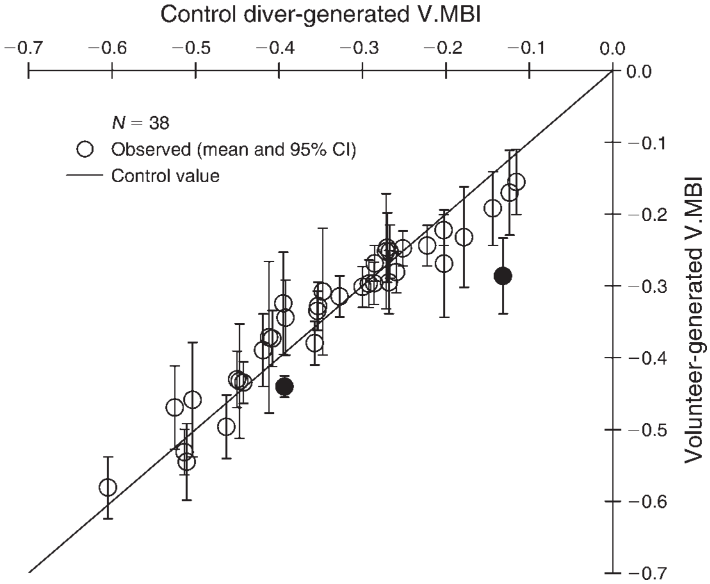


FIG. 1. Validation trials: comparison of the volunteer results with those of the control diver. The marine biodiversity index results (V.MBI) calculated from volunteers' data are compared with those calculated from the scientist control diver's data. The validation trials were performed during the four years of research from April 2002 to October 2005. Black points indicate volunteer-generated values that are significantly different from the control diver-generated values. *N* is the number of validation trials.

(CV) per team. While there was no trend in the correctness of abundance ratings with diving certification level of team members ($\rho_s = -0.097$, $N = 38$, $P = 0.562$), group size of participants ($\rho_s = -0.161$, $N = 38$, $P = 0.334$), or depth of the trial ($\rho_s = -0.302$, $N = 38$, $P = 0.065$), a negative correlation was detected with dive time of the trial ($\rho_s = -0.385$, $N = 38$, $P = 0.017$). The regression analyses, CAR (%) = –0.414[time (in minutes)] + 100.184, indicated a decrease of 4 points in CAR for every 10 minutes of dive time.

According to the α correlation test (Table 2), only two teams (5.3%) performed with an insufficient level of reliability (α, 95% CL lower bound $\leq$ 50%); three teams (7.9%) scored acceptable relationship with the control diver census (α, 95% CL lower bound $>$ 50% $\leq$ 60%), 12 teams (31.6%) scored an effective reliability level (α, 95% CL lower bound $>$ 60% $\leq$ 70%), and 21 teams (55.3%) performed from definitive to very high levels of reliability (α, 95% CL lower bound $>$ 70% $\leq$ 100%). Intra-group variation was approximately 15% (CV) per team. α correlation coefficient was not correlated with diving certification level ($\rho_s = -0.264$, $N = 38$, $P = 0.110$), group size of participants ($\rho_s = 0.070$, $N = 38$, $P = 0.675$), depth ($\rho_s = -0.131$, $N = 38$, $P = 0.433$), or dive time of the trial ($\rho_s = -0.046$, $N = 38$, $P = 0.783$), but it showed a positive trend from the first to the last year of the trials ($\rho_s = 0.711$, $N = 38$, $P < 0.001$). The regression analyses (α(%) = 6.394[time (in years)] + 62.036) indicated a 6-point increase in reliability each year.

According to the Czekanowki's proportional similarity index, SI (Table 2), 11 teams (28.9%) performed with levels of precision below the sufficiency threshold (SI, 95% CL lower bound $\leq$ 50%); 25 teams (65.8%) scored a sufficient level of precision (SI, 95% CL lower bound $>$ 50% $\leq$ 75%), and 2 teams (5.3%) scored high levels of precision (SI, 95% CL lower bound $>$ 75% $\leq$ 100%). Intra-group variation was approximately 22% (CV) per team. The similarity index was not correlated with diving certification level ($\rho_s = -0.222$, $N = 38$, $P = 0.181$), number of participants in the trial group ($\rho_s = 0.042$, $N = 38$, $P = 0.802$), depth ($\rho_s = -0.108$, $N = 38$, $P = 0.518$), or dive time of the trial ($\rho_s = 0.051$, $N = 38$, $P = 0.763$), but it showed a positive trend from the first to the last year of the trials ($\rho_s = 0.734$, $N = 38$, $P < 0.001$). The regression analyses (SI(%) = 6.923[time (in years)] + 45.687) indicated a 7-point increase in precision each year.

A comparison of V.MBI values calculated from volunteers' data with those calculated from the control diver indicated that in 36 out of 38 trials (94.7%) the volunteer generated index was not significantly different from the control diver index (Fig. 1).

Marine biodiversity monitoring

Over four years, a total of 3825 volunteer recreational divers participated in the monitoring program (Table 4). They spent a total of 13 539 hours underwater and completed 18 757 valid survey questionnaires, with a mean dive time effort per questionnaire of 43.3 minutes (95% CI 43.1–43.5; Table 4). The great majority of questionnaires (88.1%) involved rocky habitats (Table 4). The low number of useful questionnaires from sandy habitats did not allow spatiotemporal analyses of results. Conversely, for rocky habitats, most questionnaires were useful (73.8–81.2% per year).

The geographic distribution of rocky habitat surveys was homogenous over the four years (α = 0.976; $\rho_s = 0.868$; Fig. 2). Most surveys were made in the northern Tyrrhenian and Ligurian Seas, accounting for 61.9% of the total number of valid recorded questionnaires for rocky habitats. The total number of survey stations for rocky habitats was 209, of which 113 (54.1%) were surveyed for >1 year (47 stations for two years, 34 for three years, 32 for four years; detailed results from each

TABLE 4. Distribution of survey effort performed by volunteer recreational divers in the four years of research; only useful questionnaires were elaborated.

				Rocky bottom valid questionnaires		Sandy bottom valid questionnaires		Other habitat valid questionnaires	
Year	No. volunteer divers	Hours of diving	Total valid questionnaires	Recorded	Useful (%)	Recorded	Useful (%)	Recorded	Useful (%)
2002	936	2446	3342	2847	73.8	387	34.9	108	21.3
2003	1615	4459	6230	5544	79.3	428	19.2	258	46.5
2004	1214	3830	5313	4699	80.3	452	26.1	162	29.6
2005	803	2805	3872	3443	81.2	352	42.3	77	0.0
All years	3825	13 539	18 757	16 533	79.0	1619	29.9	605	31.6

Note: See *Materials and methods: Construction of the biodiversity evaluation model* for details.

survey station are available on Appendix D: Table D1). Mean depth of the surveys performed at the stations was homogeneous among years ($\alpha = 0.958$; $\rho_s = 0.898$); the most commonly surveyed depth range was between 11 and 30 m (90.0% of the stations). Also the mean time (date and hour) of the surveys performed at the stations was homogeneous among years (for the date, $\alpha = 0.851$, $\rho_s = 0.720$; for the hour, $\alpha = 0.907$, $\rho_s = 0.767$); the surveys were concentrated around the spring–summer period (83.3% of the stations had mean sampling date between May and August) and between late morning and early afternoon (84.7% of the stations had a mean sampling time between 10.00 and 15.00).

Of the 61 organismal taxa surveyed, 49.2% (30 taxa) were not common, with a sighting frequency (%SF, calculated on the total number of surveys over the four years) of $\leq$20%, 45.9% (28 taxa) were common (20% < %SF < 70%), and only 4.9% (3 taxa) were very common (%SF $\geq$ 70%; detailed data about each taxon are available on Appendix E: Table E1; taxa ranking according to sighting frequency is after Schmitt and Sullivan 1996, Darwall and Dulvy 1996). Most of the organismal taxa (54, 88.5%) had homogeneous sighting frequencies throughout the years ($\alpha = 0.925$, SE = 0.005; $\rho_s = 0.790$, SE = 0.012). Only seven taxa (11.5%) had significant annual sighting frequency differences (Fig. 3). In six cases, box crab (*Calappa granulata*), thornback ray (*Raja clavata*), John dory (*Zeus faber*), long-snouted branched seahorse (*Hippocampus ramulosus*), short-snouted seahorse (*Hippocampus hippocampus*), and flying gurnard (*Dactylopterus volitans*), the sighting frequencies had a negative trend over time (Jonckheere-Terpstra test, $P = 0.001$–0.014) and in one case, the pentagonal sea star (*Ceramaster placenta*), there were wide variations throughout the years without a trend (Jonckheere-Terpstra test, $P = 0.079$). Vegetal (H_{SHV}) and animal (H_{SHA}) biodiversity, sighting frequency of litter (%LF) and the marine biodiversity index (V.MBI) were homogeneous among years (for H_{SHV}, $\alpha = 0.868$, $\rho_s = 0.716$; for H_{SHA}, $\alpha = 0.869$, $\rho_s = 0.716$; for %LF, $\alpha = 0.939$, $\rho_s = 0.841$; for V.MBI, $\alpha = 0.826$, $\rho_s = 0.653$; Appendix E: Table E1).

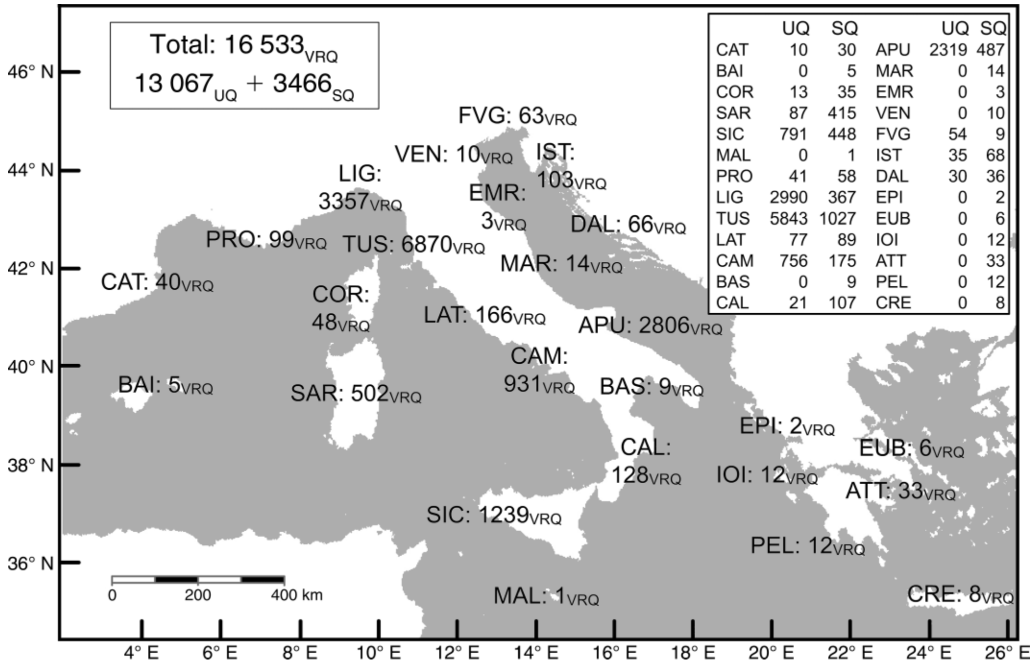
The V.MBI calculated for the 209 stations did not change significantly over the project time scale, but it had a highly significant negative correlation with latitude ($\rho_s = -0.228$, $P < 0.001$; Fig. 4). The correlation analysis performed by aggregating stations into two macro-geographic areas showed the same trend: for the western sector, stations in the Ligurian, Tyrrhenian, and Sardinian Seas, and in the Sicilian Channel gave $\rho_s = -0.231$, $P < 0.01$, N stations = 172 (the ds-10 station was excluded from the correlation analysis because it was isolated and the only one in the Gulf of Lions); for the eastern sector, stations in the Adriatic and Northern Ionian Seas gave $\rho_s = -0.294$, $P < 0.05$, N stations = 35 (the sbv-6 station was excluded from the correlation analysis because it was isolated, and the only one in the Southern Ionian Sea).

With the intention to critically evaluate the rationalization of survey effort requested to volunteers divers, the "best" match between the multivariate among-sample pattern depicted in Fig. 4, which was derived from the full assemblage of variables listed in the survey questionnaire (62: 61 organismal taxa plus litter), and that from random subsets of the variables was determined. The best explanatory variables which generated the same multivariate sample pattern as the full list, turned out to be the subset of 27 organismal taxa listed in Table 3, representing the 43.5% of the original list of variables.

DISCUSSION

Validation trials: quality of recreational volunteer-generated data

The levels of accuracy performed during validation trials were encouraging given the number of species surveyed and the recreational dive profile (i.e., the divers did not follow pre-oriented transects, but they dived following the normal recreational dive path for a given dive site). Accuracy was comparable to that performed by conservation volunteer divers on precise transects in other projects (Mumby et al. 1995, Darwall and Dulvy 1996), or in community-based terrestrial monitoring (Evans et al. 2000). At greater than the high level of accuracy of 80% (categorized high by Delaney et al. 2008), the accuracy reached by volunteers in some trials was particularly impressive, as impressive was the results that only in two trials out of 38 (5.3%), the V.MBI



	UQ	SQ		UQ	SQ
CAT	10	30	APU	2319	487
BAI	0	5	MAR	0	14
COR	13	35	EMR	0	3
SAR	87	415	VEN	0	10
SIC	791	448	FVG	54	9
MAL	0	1	IST	35	68
PRO	41	58	DAL	30	36
LIG	2990	367	EPI	0	2
TUS	5843	1027	EUB	0	6
LAT	77	89	IOI	0	12
CAM	756	175	ATT	0	33
BAS	0	9	PEL	0	12
CAL	21	107	CRE	0	8

FIG. 2. Geographic distribution of the survey effort performed on rocky bottom habitats over the four years of research (2002–2005). The total number of valid recorded questionnaires (VRQ) was divided into useful questionnaires (UQ), those coming from survey stations, and sparse questionnaires (SQ), those coming from diving sites that failed to reach an annual quorum of 10 recorded questionnaires. Key to site abbreviations: APU, Apulia; ATT, Attica; BAI, Balearic Islands; BAS, Basilicata; CAL, Calabria; CAM, Campania; CAT, Catalonia; COR, Corsica; CRE, Crete; DAL, Dalmatia; EMR, Emilia-Romagna; EPI, Epirus; EUB, Euboea; FVG, Friuli-Venezia Giulia; IOI, Ionian Islands; IST, Istria; LAT, Latium; LIG, Liguria; MAL, Malta; MAR, Marches; PEL, Peloponnesus; PRO, Provence; SAR, Sardinia; SIC, Sicily; TUS, Tuscany; VEN, Veneto.

generated by the volunteers was significantly different from the control value generated by the scientist diver.

Since temporal and spatial comparisons of sites are based upon the survey data obtained by volunteers, attaining high consistency is, therefore, essential for comparative data analyses. The level of consistency reached by volunteers during validation trials is comparable to that performed by conservation volunteer divers on precise transects (Mumby et al. 1995), and in some trials consistency resulted greater than 70%.

One trend related to both data accuracy and consistency emerged, with the presence of a clear improvement in data quality from the first through to the last year of validation trials. This result was not surprising, considering the key presence of positive feedback during the survey program. Feedback, corrections and learning were given by trained professional divers (trained divemasters and instructors that guided the volunteers during the dive) under normal survey conditions. After each dive, trained professional divers debriefed volunteer divers to highlight areas of weakness, source of inaccuracy, and taxa misidentification. Among the several potential sources of group variation, diligence may explain the negative correlation between the level of consistency reached in the validation trial and the diving certification level of group members. First level divers tend to stay in pairs, close to each other, and to follow the divemaster along the dive path with attention; in contrast more highly qualified divers are less diligent, and tend to diversify from the path, consequently recording different sightings and leading to decreased correlation among recorded data.

Similarly to conservation volunteers on precise transects (Mumby et al. 1995, Bell 2007), the positive correlation between correct identification and the taxa presence frequency in the validation trials indicated that recreational volunteers were more accurate in recording the most frequent/straightforward taxa, while they were less accurate with rare/cryptic taxa, even if the identification of these of taxa was specifically addressed in the training program. The intercept of the regression analyses between correct identification and taxa presence frequency suggested that even the rarest taxa tend to be correctly identified by more than 50% of volunteers, which represents sufficient correct identification.

The negative regression between dive time and the capability of volunteers to assign precise ordinal abundance ratings indicates that after 45 minutes of dive, which represent a mean recreational dive time in temperate water, the correctness of abundance ratings is still above 80%, and that after 60 minutes (long recreational dive time in temperate water) the correctness of abundance ratings is still 75%. These data

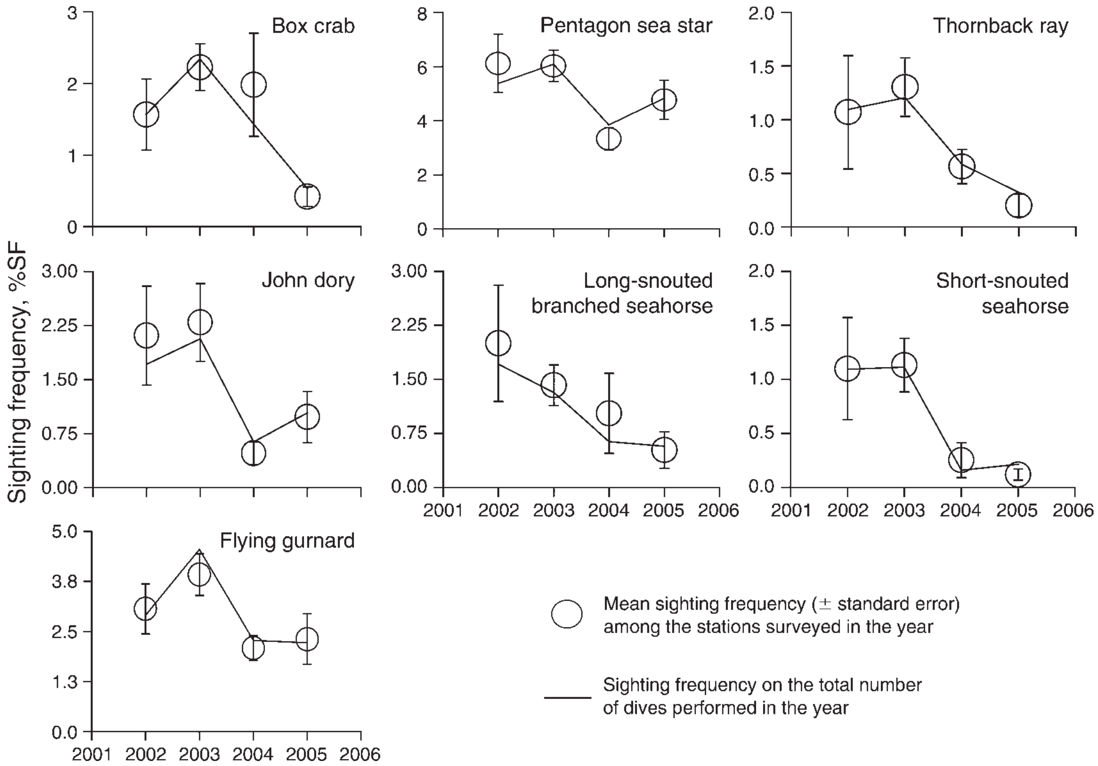


FIG. 3. Taxa with non-homogeneous sighting frequencies among the years: box crab (*Calappa granulata*), pentagon sea star (*Ceramaster placenta*), thornback ray (*Raja clavata*), John dory (*Zeus faber*), long-snouted branched seahorse (*Hippocampus ramulosus*), short-snouted seahorse (*Hippocampus hippocampus*), flying gurnard (*Dactylopterus volitans*). For these taxa the sighting frequency (%SF, percentage of dives where the taxon was sighted) is represented over the four-year study.

suggest that only after very long dive times (which are highly improbable for recreational dives in temperate waters) physical, physiological, and psychological factors (tiredness, chilling, possible nitrogen narcosis effects, anxiety, memory recall, fatigue) can significantly reduce survey performance at the depths were recreational volunteers performed (4–28 m) and with a normal recreational SCUBA gear.

Problems and limitations

Some studies show that under conditions of appropriate recruitment and training, volunteer-collected data are qualitatively equivalent to those collected by professional researchers and useful for resource management (Darwall and Dulvy 1996, Schmitt and Sullivan 1996, Fore et al. 2001, Greenwood 2003, Newman et al. 2003, Pattengill-Semmens and Semmens 2003, Boudreau and Yan 2004, Bell 2007, Tobias and Brightsmith 2007). There were a number of features of this study that indicated reliability of the volunteer-collected data presented here. The points that showed that acceptable level of reliability was achieved are outlined below:

1) The data were markedly consistent across years, indicating a strong degree of reliability, as in our previous volunteer-based marine conservation monitoring project (Goffredo et al. 2004);

2) Trends in this data set were corroborated by data in scientific literature and databases;

3) The results of the validation trials indicated that volunteers performed with levels of accuracy and consistency comparable to those of conservation volunteers on precise transects in other projects (Mumby et al. 1995, Darwall and Dulvy 1996, Evans et al. 2000).

The reasons why reliability was achieved are:

1) Volunteers were trained and assisted during data collection in the field by dive guides and instructors who had previously attended workshops and received training on project objectives and methodology by professional researchers;

2) The method was designed to be suitable for amateurs (i.e., user-friendly questionnaire and taxa that are easily recognizable by recreational divers);

3) Information requested on the questionnaire such as dive location, depth, dive time, and habitat are details that most divers routinely record in their personal dive logs, whether the purpose of the dive is recreational or for data collection; selection of appropriate tasks for volunteers at the research planning stage of the project is fundamental, since volunteer skills and abilities vary, and we only wanted volunteers collect data for which they could be trained quickly and reliably.

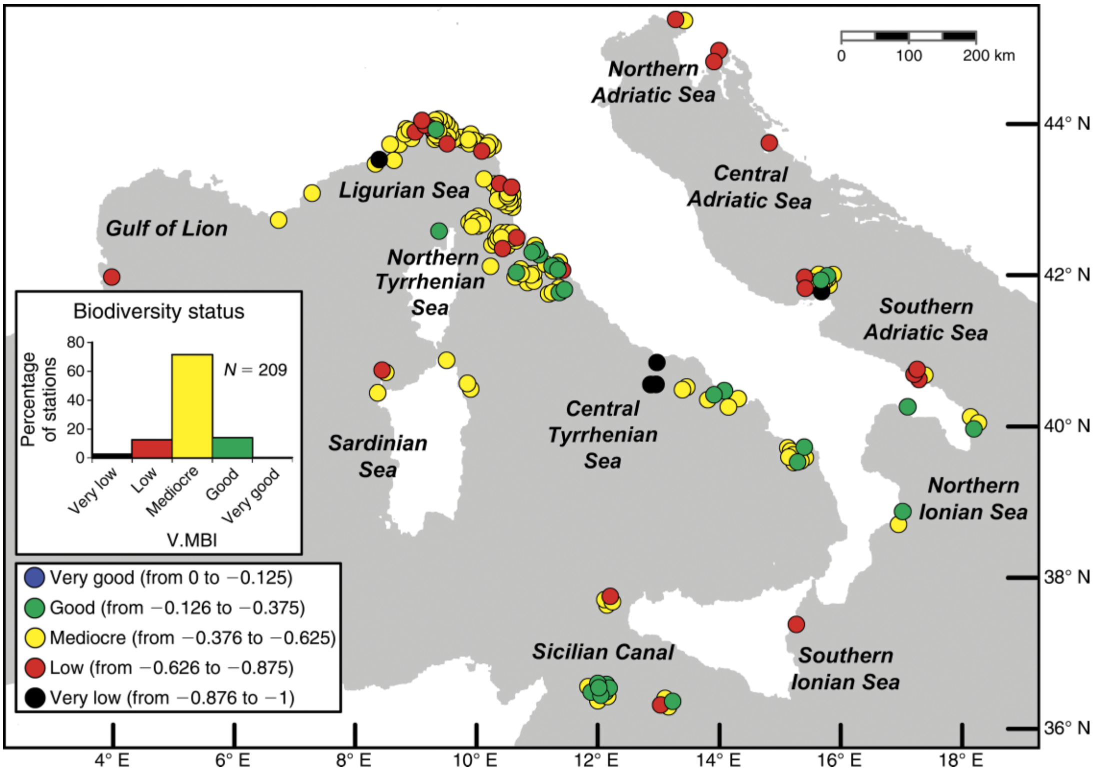


FIG. 4. Marine biodiversity index (V.MBI) in the 209 stations surveyed in the four years of research (2002–2005). Marine biodiversity measured by the index in the 209 stations gave a Gaussian distribution, with most stations (71.3%) being of mediocre status. The index did not show maximum status class (very good) in any of the stations. Summary measures by region are presented in Appendix F: Figs. F1–F3.

The primary limiting factor in involving citizen-volunteers was the difficulty in obtaining data homogeneously spatially distributed. In fact, most questionnaires came from rocky habitats along Ligurian and northern Tyrrhenian sea coasts. This biased sampling effort may be explained by recreational divers' preference for rocky habitats, which tend to be more biodiverse and are therefore more interesting to visit than sedimentary habitats (Goffredo et al. 2004). Attempts made to encourage data collection on sandy bottoms (in the form of prizes; as in Goffredo et al. 2004) were not successful in increasing surveys in this habitat. The northwestern coast was surveyed more because: (1) the Tyrrhenian and Ligurian Seas are more attractive to divers because of water clarity compared to the central northern Adriatic Sea; (2) there are proportionately more diving centers along the Northern Tyrrhenian and Ligurian sea coasts providing logistical appeal (21.4 centers per 100 km of coast vs. a national average of 6.7); (3) the national headquarters of some of the diving agencies that officially supported the project are located in northern Italy.

Bathymetric and temporal survey distribution reflected the typical pattern of recreational diver activity. Normally, international diving school agencies recommend 30 m as the maximum depth (World Recreational Scuba Training Council 2006) and the preferred period for diving is the warm season during the daytime (only advanced divers perform night dives).

Volunteer participation

Participation reached its peak in the second year when the popular national scientific magazine, Quark, and the Italian Tour Operators Association became official partners. They helped to promote the project and offered prizes to reward volunteers. After the second year, there was a drop in the number of participants, especially in the fourth year (–33.9% in 2005 compared to the previous year). This drop may have been due to the departure of one of the partner diving agencies from the project, poor weather during the summer of 2005 and, according to interviews with tour operators, the economic crisis that limited general public expenditure on recreation. Unfortunately we did not collect data on the "enjoyment" of the survey dives compared to non-survey dives as experienced by the divers. However, the mean annual survey effort per individual volunteer constantly increased over the four-year period (mean number of questionnaires recorded /hours of diving a year per volunteer: first year 3.6/2.6, second year 3.9/2.8, third year 4.4/3.2, fourth year 4.8/3.5). This positive trend may reflect the growing interest and loyalty of volunteer divers to the project.

Assessed biodiversity and environmental conditions

Given that our study lasted only four years, it is not surprising that sighting frequencies of most taxa were consistent over the years. Of the seven exceptions, six showed significant declines. It is known that four of these have declined in the long term in the Mediterranean sea due to over-fishing or habitat damage (thornback ray, John Dory, and the two seahorses: Garofalo et al. 2003, Boudouresque 2004, Vrgoč et al. 2006).

The fact that the presence of litter in the environment did not substantially change over a four-year period is also expected, unless clean-up operations are performed (Davenport and Davenport 2006).

Our findings regarding increasing of the V.MBI with decreasing latitude can be interpreted as an improvement in environmental conditions at coastal stations going from north to south. An alternative explanation is that the detected variation is just a latitudinal variation, given the geographic scale. The first interpretation is supported by the data from the Italian Ministry of the Environment. Concurrent with this study, the Ministry conducted sea water quality surveys, including parameters reflecting hygiene/health risks (CAM index, sea water classification; Italian Ministry of the Environment and Land and Sea Protection 2006). For areas overlapping with those monitored by our study, data from the Italian Ministry corroborate negative correlations between latitude and environmental quality: for the western region $\rho_s = -0.277$, $P < 0.01$, 114 stations; for the eastern region $\rho_s = -0.543$, $P < 0.001$, 46 stations. In the seas surrounding Italy stressors (overexploitation of fisheries, eutrophication, domestic waste, hydrocarbons and oil, heavy metals,) are more prominent in the northern areas than in southern ones with some northern locations extremely degraded (Caddy 1998, Danovaro 2003, Thibaut et al. 2005). In the northern parts of the Western Mediterranean, a marked reduction in overall marine biodiversity has also resulted from both biological invasions of alien species and the largest mass mortality event of benthic invertebrates ever recorded in the Mediterranean basin, which was most probably caused by climatic anomalies (Boudouresque and Verlaque 2005, Linares et al. 2005).

According to the BEST test of searching over subsets of variables for a combination that optimizes the survey effort, 27 out of 62 taxa (43.5% of the original assemblage) were sufficient to generate the same multivariate sample pattern. For future monitoring research, limitation of items to the most necessary could, one hand lead to a reduction in effort during both volunteer training and field work, but on the other hand, it could limit the appeal of the project to potential volunteers. Removing attractive species from the questionnaire (for example red coral, yellow cluster anemone, dotted sea slug, common octopus, lobster, spider crab, moray eel, sea raven, rainbow wrasse, anglerfish) is likely to have decreased volunteers' enjoinment and loyalty, and also the educational potential of the project. Adding charismatic organisms that citizen volunteers are likely to see to the survey in order to give them something to report with satisfaction is an approach successfully experimented in ornithological studies (Greenwood 2007).

CONCLUSIONS

This project successfully involved citizens that use the sea for recreational purposes (such as tourist divers and snorkelers) in the collection of data recording the presence of biological taxa and litter. The conclusions that can be drawn from this work are:

1) Trained recreational divers achieve an acceptable level of accuracy and consistency.

2) Recreational diver-based surveys can provide useful information in marine biodiversity surveys, significantly reducing financial and time costs. With the participation of recreational divers we were able to amass a large data set, covering a wide geographic area, over a relatively short period of time. We estimated that in order to collect the same amount of data obtained by the volunteers in this study a single professional would have needed 45 years and more than US$4 758 000.

3) Recreational divers tend to concentrate on rocky bottoms, in a scheme where they were not forced to cover any habitats in particular.

4) The quality of data improved with time, as the survey organizers and instructors gained experience of how to brief volunteers.

5) The consistency of the records of high level divers was less than the consistency of low-level divers.

6) A subset of the taxa would have been adequate for the survey purposes, though it was probably useful to include at least some of the "unnecessary" taxa in order to maintain the interest of the volunteers.

In our experience, and of other institutes (Darwall and Dulvy 1996, U.S. Environmental Protection Agency 1997, Evans et al. 2000, Foster-Smith and Evans 2003, Bhattacharjee 2005, Sharpe and Conrad 2006, Bell 2007), "citizen science" can complement and augment conventional methods, and it can be a key solution to personnel needed to carry out research. Given the scarce government resources, the role of citizens and the civil community in monitoring is especially important, even when volunteers need special skills, as those necessary for exploring the underwater environment.

Citizen involvement as ecological research operators improves scientific literacy and environmental awareness and education amongst all age groups in the community (Evans et al. 2005), and determines a more sustainable approach to the environment (Medio et al. 1997). Environmental education provides the long-term solution to sustainable management of the environment. However, formal education operates under severe curriculum constraints and has been at best only partially successful in achieving this goal (Holdren and Ehrlich 1971, Evans 1988). There is a need therefore for new educational initiatives. "Divers for the Environment: Mediterranean Underwater Biodiversity Project" was one of such initiative. Education, the "citizen science" approach, the development of an interdisciplinary mentality in researchers, and the realization of research projects that take into account the needs and motivations of people are practical efforts necessary to complete the mission of modern conservation biology (Meffe et al. 2006). This report may inspire other researchers to incorporate citizen scientists in their projects.

ACKNOWLEDGMENTS

We dedicate this study to Umberto Pepoli, the best diver educator, and the first to strongly believe in this project, who left us on 1 May 2007. Special thanks go to all the divers who, by participating with enthusiasm in the collection of data, have made this study possible. Sources of funding were the Italian Government (Ministry of the University and Research, Ministry of the Environment and Land and Sea Protection), the ASTOI (Association of Italian Tour Operators), the Association of Diving agencies operating in Italy (IDEA, International Diving Educators Association; PADI, Professional Association of Diving Instructors; PSS, Professional Scuba Schools; SNSI, Scuba Nitrox Safety International; SSI, Scuba Schools International), the popular scientific magazine Quark (Hachette-Rusconi), the environmental association Underwater Life Project, and the Project Aware Foundation. G. Neto and F. Turano took the photos of the survey questionnaire. Marine and Freshwater Science Group Association realized educational programs connected to the project in the public schools. E. Pignotti (Statistical Analysis Task Force, Rizzoli Orthopaedic Institute Research Center of Bologna) gave valuable advice on statistical analysis of data. J. J. D. Greenwood (British Trust for Ornithology), O. Langmead (University of Plymouth), D. Medio (Halcrow Group Ltd), C. Lewis (Florida Fish and Wildlife Commission), M. Mukherjee (University of Georgia), J. Bilewitch (State University of New York), E. L. McLean (University of Puerto Rico), N. Kirk (Auburn University), E. Caroselli (University of Bologna), G. Meffe (University of Florida), and D. Couvet (Muséum National d'Histoire Naturelle, France) gave advice which improved the manuscript. P. Addis, G. Boscoscuro, A. Corti, and R. Martinez personally contributed to the realization of this project. Special thanks go to the Egyptian Tourist Authority, Ministry of Tourism of the Arab Republic of Egypt, for its current support to citizen-science projects developed in the Red Sea ⟨www.STEproject.org⟩. The comments of two anonymous reviewers are gratefully acknowledged.

LITERATURE CITED

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

Balmford, A., et al. 2005. The Convention on Biological Diversity's 2010 target. Science 307:212–213.

Bell, J. 2007. The use of volunteers for conducting sponge biodiversity assessment and monitoring using a morphological approach on Indo-Pacific coral reefs. Aquatic Conservation: Marine and Freshwater Ecosystems 17:133–145.

Bhattacharjee, Y. 2005. Ornithology. Citizen scientists supplement work of Cornell researchers. Science 308:1402–1403.

Boudouresque, C. F. 2004. Marine biodiversity in the Mediterranean: status of species, populations and communities. Scientific Reports Port-Cros National Park 20:97–146.

Boudouresque, C. F., and M. Verlaque. 2005. Nature conservation, Marine Protected Areas, sustainable development and the flow of invasive species to the Mediterranean Sea. Scientific Reports Port-Cros National Park 21:29–54.

Boudreau, S., and N. Yan. 2004. Auditing the accuracy of a volunteer-based surveillance program for an aquatic invader *Bythotrephes*. Environmental Monitoring and Assessment 91: 17–26.

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

Bruyere, B., and S. Rappe. 2007. Identifying the motivations of environmental volunteers. Journal of Environmental Planning and Management 50:503–516.

Caddy, J. F. 1998. Issues in Mediterranean fisheries management: geographical units and effort control. Studies and Reviews of General Fisheries Council for the Mediterranean Number 70. FAO, Rome, Italy. ⟨ftp://ftp.fao.org/docrep/fao/006/w9941e/w9941e00.pdf⟩

Campbell, L. M., and C. Smith. 2006. What makes them pay? Values of volunteer tourists working for sea turtle conservation. Environmental Management 38:84–98.

Clarke, K. R., P. J. Somerfield, and R. N. Gorley. 2008. Testing of null hypotheses in exploratory community analyses: similarity profiles and biota-environment linkage. Journal of Experimental Marine Biology and Ecology 366:56–69.

Coakes, S. J., and L. G. Steed. 1997. SPSS: analysis without anguish (version 6.1 for IBM and Macintosh users). Jacaranda Wiley, Milton, Queensland, Australia.

Cohn, J. P. 2008. Citizen science: can volunteers do real research? BioScience 58:192–197.

Danovaro, R. 2003. Pollution threats in the Mediterranean Sea: an overview. Chemistry and Ecology 19:15–32.

Darwall, W. R. T., and N. T. 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.

Davenport, J., and J. L. Davenport. 2006. The impact of tourism and personal leisure transport on coastal environments: a review. Estuarine Coastal and Shelf Science 67:280–292.

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.

Douglas, S., and A. Lawrence. 2004. Tropical biologist, local people and conservation: new opportunities for collaboration. Trends in Ecology and Evolution 19:634–638.

Evans, C., E. Abrams, R. Reitsma, K. Roux, L. Salmonsen, and P. P. Marra. 2005. The Neighborhood Nestwatch Program: participant outcomes of a citizen-science ecological research project. Conservation Biology 19:589–594.

Evans, S. M. 1988. Man and the environment: the need for a more realistic approach to teaching ecology. Journal of Biological Education 22:136–138.

Evans, S. M., and A. C. Birchenough. 2001. Community-based management of the environment: lessons from the past and options for the future. Aquatic Conservation: Marine and Freshwater Ecosystems 11:137–147.

Evans, S. M., A. C. Birchenough, and H. Fletcher. 2000. The value and validity of community-based research: TBT contamination of the North Sea. Marine Pollution Bulletin 40:220–225.

Flynn, B. B., R. G. Schroeder, and S. Sakakibara. 1994. A framework for quality management research and an associated measurement instrument. Journal of Operations Management 11:339–366.

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

Foster-Smith, J. 2000. The Northumberland Coast Marine Biodiversity Project: building records of common species. Porcupine Marine Natural History Society Newsletter 4:3–13.

Foster-Smith, J., and S. M. Evans. 2003. The value of marine ecological data collected by volunteers. Biological Conservation 113:199–213.

Garofalo, G., M. Gristina, F. Fiorentino, F. Cigala Fulgosi, G. Norrito, and G. Sinacori. 2003. Distributional pattern of rays (Pisces, Rajidae) in the Strait of Sicily in relation to fishing pressure. Hydrobiologia 503:245–250.

Garrod, B., and S. Gössling. 2008. Introduction. Pages 3–28 *in* B. Garrod and S. Gossling, editors. New frontiers in marine tourism: diving experience, sustainability, management. Elsevier, Amsterdam, The Netherlands.

Gilmour, J., and D. A. Saunders. 1995. Earthwatch: an international network in support of research on nature conservation. Pages 627–633 *in* D. A. Saunders, J. L. Craig, and E. M. Mattiske, editors. Nature conservation 4: the role of networks. Surrey Beatty and Sons, Chipping Norton, New South Wales, Australia.

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

Greenwood, J. J. D. 2003. The monitoring of British breeding birds: a success story for conservation science? Science of the Total Environment 310:221–230.

Greenwood, J. J. D. 2007. Citizens, science and bird conservation. Journal of Ornithology 148:S77–S124.

Grime, J. P. 1997. Biodiversity and ecosystem function: the debate deepens. Science 277:1260–1261.

Hair, J. F., R. E. Anderson, R. L. Tatham, and W. C. Black. 1995. Multivariate data analysis. Prentice-Hall International, Englewood Cliffs, New Jersey, USA.

Hames, R. S., K. V. Rosenberg, J. D. Lowe, S. E. Barker, and A. A. Dhondt. 2002. Adverse effects of acid rain on the distribution of the wood thrush *Hylocichla mustelina* in North America. Proceedings of the National Academy of Sciences USA 99:11235–11240.

Harrison, C. M., J. Burgess, and J. Clark. 1998. Discounted knowledge: farmers' and residents' understandings of nature conservation goals and policies. Journal of Environmental Management 54:304–320.

Hodgson, G. 1999. A global coral reef assessment. Marine Pollution Bulletin 38:345–355.

Holdren, J. P., and P. R. Ehrlich. 1971. Global ecology: readings towards a rational strategy for man. Harcourt Brace Jovanovich, New York, New York, USA.

Hughey, K. F. D., R. Cullen, G. N. Kerr, and A. J. Cook. 2004. Application of the pressure-state-response framework to perceptions reporting of the state of the New Zealand environment. Journal of Environmental Management 70:85–93.

Italian Ministry of the Environment and Land and Sea Protection. 2006. Si.Di.Mar. Protection of the sea: monitoring program of coastal marine environment. Si.Di.Mar., Rome, Italy.

Kendall, M. A., and J. R. Lewis. 1986. Temporal and spatial patterns in the recruitment of *Gibbula umbilicalis*. Hydrobiologia 142:15–22.

Kovács, A., U. C. C. Mammen, and C. V. Wernham. 2008. European monitoring for raptors and owls: state of the art and future needs. Ambio 37:408–412.

Lambert, J. D., K. P. McFarland, C. C. Rimmer, S. D. Faccio, and J. L. Atwood. 2005. A practical model of Bicknell's Thrush distribution in the Northeastern United States. Wilson Bulletin 117:1–11.

Linares, C., R. Coma, D. Díaz, M. Zabala, B. Hereu, and L. Dantart. 2005. Immediate and delayed effects of a mass mortality event on gorgonian population dynamics and benthic community structure in the NW Mediterranean. Marine Ecology Progress Series 305:127–137.

Lodge, D. M., S. Williams, H. J. MacIsaac, K. R. Hayes, B. Leung, S. Reichard, R. N. Mack, P. B. Moyle, M. Smith, D. A. Andow, J. T. Carlton, and A. McMichael. 2006. Biological invasions: recommendations for U.S. policy and management. Ecological Applications 16:2035–2054.

Lohrer, A. M., S. F. Thrush, and M. M. Gibbs. 2004. Bioturbators enhance ecosystem function through complex biogeochemical interactions. Nature 431:1092–1095.

Magurran, A. E. 1988. Ecological diversity and its measurements. Princeton University Press, Princeton, New Jersey, USA.

Medio, D., R. F. G. Ormond, and M. Pearson. 1997. Effects of briefings on rates of damage to corals by scuba divers. Biological Conservation 79:91–95.

Meffe, G. K., D. Ehrenfeld, and R. F. Noss. 2006. Conservation biology at twenty. Conservation Biology 20:595–596.

Mumby, P. J., A. R. Harborne, P. P. Raines, and J. M. Ridley. 1995. A critical assessment of data derived from Coral Cay Conservation volunteers. Bulletin of Marine Science 56:737–751.

Myers, J. H., D. Simberloff, A. M. Kuris, and J. R. Carey. 2000. Eradication revisited: dealing with exotic species. Trends in Ecology and Evolution 15:316–320.

National Audubon Society. 2006. The Christmas bird count historical results. National Audubon Society, New York, New York, USA. ⟨http://www.audubon.org/bird/cbc/hr/index.html⟩

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.

Nunnally, J. C. 1967. Psychometric theory. First edition. McGraw-Hill, New York, New York, USA.

Oberhauser, K., I. Gebhard, and C. Cameron. 2007. Parasitism of Monarch Butterflies (*Danaus plexippus*) by *Lespesia archippivora* (Diptera: Tachinidae). American Midland Naturalist 157:312–328.

Oliver, I., and A. J. Beattie. 1993. A possible method for rapid assessment of Biodiversity. Conservation Biology 7:562–568.

Pattengill-Semmens, C. V., and B. X. Semmens. 2003. Conservation and management applications of the reef volunteer fish monitoring program. Environmental Monitoring and Assessment 81:43–50.

Pearson, D. L. 1994. Selecting indicator taxa for the quantitative assessment of biodiversity. Philosophical Transactions: Biological Sciences 345:75–79.

Peterson, R. A. 1994. A meta-analysis of Cronbach's coefficient alpha. Journal of Consumer Research 21:381–391.

Ryan, R. L., R. Kaplan, and R. E. Grese. 2001. Predicting volunteer commitment in environmental stewardship programmes. Journal of Environmental Planning and Management 44:629–648.

Sale, P. F., and W. A. Douglas. 1981. Precision and accuracy of visual census technique for fish assemblages on patch reefs. Environmental Biology of Fishes 6:333–339.

Schmitt, E. F., and K. M. Sullivan. 1996. Analysis of a volunteer method for collecting fish presence and abundance data in the Florida Keys. Bulletin of Marine Science 59:404–416.

Sharpe, A., and C. Conrad. 2006. Community based ecological monitoring in Nova Scotia: challenges and opportunities. Environmental Monitoring and Assessment 113:395–409.

Sheil, D., and A. Lawrence. 2004. Tropical biologist, local people and conservation: new opportunities for collaboration. Trends in Ecology and Evolution 19:634–638.

Stokes, P., M. Havas, and T. Brydges. 1990. Public participation and volunteer help in monitoring programs; an assessment. Environmental Monitoring and Assessment 15: 225–229.

Therriault, T. W., and J. Kolasa. 2000. Explicit links among physical stress, habitat heterogeneity and biodiversity. Oikos 89:387–391.

Thibaut, T., S. Pinedo, X. Torras, and E. Ballesteros. 2005. Long-term decline of the populations of Fucales (*Cystoseira* spp. and *Sargassum* spp.) in the Alberes coast (France, North-western Mediterranean). Marine Pollution Bulletin 50: 1472–1489.

Tobias, J. A., and D. J. Brightsmith. 2007. Distribution, ecology and conservation status of the Blue-headed Macaw *Primolius couloni*. Biological Conservation 139:126–138.

Trumbull, D. J., R. Bonney, D. Bascom, and A. Cabral. 2000. Thinking scientifically during participation in a citizen-science project. Science Education 84:265–275.

U.S. Environmental Protection Agency. 1997. What is volunteer monitoring? U.S. Environmental Protection Agency, Washington, D.C., USA. ⟨http://epa.gov/owow/monitoring/volunteer/⟩

Vrgoč, N., S. Krstulović Šifner, V. Dadić, and S. Jukić-Peladić. 2006. Demographic structure and distribution of John Dory, *Zeus faber* L. 1758, in the Adriatic Sea. Journal of Applied Ichthyology 22:205–208.

World Recreational Scuba Training Council. 2006. Dive standards and medical statement. World Recreational Scuba Training Council, Jacksonville, Florida, USA. ⟨http://www.wrstc.com/downloads.php⟩

APPENDIX A

Survey questionnaires: section with photographs to identify the surveyed taxa (*Ecological Archives* A020-081-A1).

APPENDIX B

Survey questionnaires: section with a form to record data (*Ecological Archives* A020-081-A2).

APPENDIX C

Volunteer divers training and involvement (*Ecological Archives* A020-081-A3).

APPENDIX D

Geographic coordinates, number of useful questionnaires, bathymetry of survey, and moment of survey in the survey stations on rocky bottom (*Ecological Archives* A020-081-A4).

APPENDIX E

Taxa sighting frequency, observed biodiversity, vegetal and animal litter sighting frequency, and marine biodiversity index in the survey stations on rocky bottom (*Ecological Archives* A020-081-A5).

APPENDIX F

Summary measures by region of marine biodiversity index (V.MBI) in the four years of research (2002–2005) (*Ecological Archives* A020-081-A6).