Reliability and utility of citizen science reef monitoring data collected by Reef Check Australia, 2002 - 2015

Terence Done¹,²,⁹*, Chris Roelfsema²,³, Andrew Harvey⁴,⁵, Laura Fantozzi⁶, Jocelyn Hill⁷, Marie-Lise Schläppy⁸, Alexandra Lea², Anne Bauer-Civiello², and Jennifer Loder²

¹ Australian Institute of Marine Science, PMB #3, Townsville MC, QLD, 4810, Australia
² Reef Check Australia, PO Box 13204 George St, Brisbane, QLD 4003 Australia
³ Remote Sensing Research Centre, School of Geography, Planning and Environmental Management, University of Queensland, Brisbane QLD 4072, Australia
⁴ Faculty of Medicine and Biomedical Sciences, University of Queensland, Brisbane QLD 4072, Australia
⁵ MantaWatch, 20-22 Wenlock Road, London N1 7GU, United Kingdom
⁶ Scripps Institution of Oceanography, UC San Diego, La Jolla, California 92037 USA
⁷ Coral Reef Alliance, 1330 Broadway, Suite 1602, Oakland, California 94612, USA
⁸ Environmental Research Institute, University of Highlands and Islands, Thurso, UK
⁹ Museum of Tropical Queensland, 78-102 Flinders St, Townsville, Queensland, 4810, Australia

*Corresponding Author: t.done@aims.gov.au
Abstract

Reef Check Australia (RCA) has collected data on benthic composition and cover at > 70 sites along > 1000 km of Australia’s Queensland coast from 2002 – 2015. This paper quantifies the accuracy, precision and power of RCA benthic composition data, to guide its application and interpretation. A simulation study established that the inherent accuracy of the Reef Check point sampling protocol is high (< ± 7% error absolute), in the range of estimates of benthic cover from 1% to 50%. A field study at three reef sites indicated that, despite minor observer- and deployment-related biases, the protocol does reliably document moderate ecological changes in coral communities. The error analyses were then used to guide the interpretation of inter-annual variability and long term trends at three study sites and assessments in RCA’s major 2002-2015 data series for the Queensland coast.

Key words

citizen science; benthos; coral monitoring; Great Barrier Reef, subtropical reefs
Introduction

Long-term ecological monitoring makes important contributions to the science, adaptive management and public perception of the state of coral reefs. ‘Citizen scientists’ have long been notable contributors to long-term reef monitoring (Hodgson 1999, 2001; Beeden et al. 2014; Loder et al. 2015). Along the Queensland coastline, including the Great Barrier Reef (GBR), major citizen science initiatives include Eye on the Reef (Beeden et al. 2014), Coralwatch (Marshall et al. 2014) and Reef Check Australia (RCA) (Loder et al. 2015), all with a major focus on recreational dive sites.

RCA (established in 2001) is a not for profit, registered charity with a small staff, a large volunteer membership base and established governance framework. Program objectives are organized into three main streams; conservation; education and citizen science (Reef Check Australia 2013). In its citizen science stream (the focus of this paper), RCA trains and coordinates SCUBA and snorkel volunteers to contribute data sets on benthos, substratum, invertebrates, fish and human impacts on reefs. At more than 70 sites along the Queensland coast, volunteers annually visit long-term monitoring sites to record data. Critical to the success and sustainability of citizen science projects such as this are opportunities for the meaningful engagement of volunteers and the collection of data of suitable quality for the research question (Dickinson et al. 2010). RCA has a strong record of engagement. Since 2002, 218 RCA volunteers have monitored > 60 sites in the GBR and, since 2009, an additional 15 in South East Queensland (SEQ) (Fig. 1). Volunteers also work with staff to share data through multiple formats, including scientific journals (e.g. Roelfsema et al. 2016), reports for program stakeholders and community members (e.g. Welch et al. 2016), social media, community talks, infographics, and community events. In relation to data quality, RCA’s objective is to produce benthic data that complements formal government monitoring programs in the provision of indicators of ecological conditions at a spatial scale appropriate for site management and condition reporting.

RCA’s methods are derived from those of Reef Check (RC), the international volunteer, community-based reef monitoring program (Hodgson 1999) that aspires to ‘provide a valuable method to detect broad-brush changes on a local, regional and global scale, as well as increasing public support for coral reef conservation’ (Hodgson 1999, p 345). RC affiliates around the world use a standard point intercept transect (PIT) sampling protocol and report on percent cover of 10 standard benthic categories: live hard coral, recently dead hard coral, soft coral, fleshy seaweed, sponge, other
benthos, rock, rubble, sand, and silt/clay. A principal focus of most reef monitoring is live hard coral, because of its functional roles as key builder of reef structure and key provider of complex, rigid habitat for much other reef biodiversity (McClanahan et al. 2002). Marine pollution, particularly in the form of agricultural runoff, can be a key driver of phase-shifts from coral to non-coral benthic dominance (Done 1992), so it is important that RC’s protocols should reliably document both catastrophic change and long-term trends in live hard coral cover, and transformation to alternate states such as soft corals, fleshy algae or sponges. If RCA’s field estimates of benthic cover are to constitute a useful basis for such scientific inference or management decision making, the degree of uncertainty associated with such estimates needs to be known. On the GBR, episodic losses in live hard coral generally range from 12 to 43% (caused by storms, coral predators, diseases and bleaching events; Osborne et al. 2011), and incremental hard coral gains over 5 – 10 years in recovering shallow slope sites have been reported at around 4% per year (Done et al. 2010; Osborne et al. 2011). Therefore, monitoring to track changes in cover must be able to document changes of these magnitudes.

Accordingly, this paper investigates the reliability and utility of the RCA benthic monitoring data. As with many citizen science projects, this is especially important given that surveyors may change from one year to the next, and that, while study sites are well defined, transect lines within study sites are not permanently marked. Here, we quantify these effects statistically through replicated field studies at representative sites. Our statistical null hypothesis is that there is no difference in estimates of percent cover among observers using the same transect line, or using different placements of the transect line (henceforth referred to as ‘deployments’) within the delineated study area. We consider that, despite any such observer or deployment effects, an effective citizen-based long-term monitoring program still needs to be able to reliably distinguish real change from sampling noise in key indicators such as the percentage cover of macro algae and hard coral.

Methods

Routine benthic monitoring

RCA benthic monitoring is carried out only by volunteers who have undertaken standardised training and demonstrated competency in identification of benthic categories through both a photographic identification exam (85% pass rate) and in-situ identification skills (95% pass rate). In routine
monitoring, RCA volunteers record corals, algae and sponges down to the level of specific growth forms (25 categories) but for the purposes of the present study records are grouped into 10 standard RC categories. Data may be accessed through the RCA web site http://www.reefcheckaustralia.org.

The term ‘site’ refers here to a narrow long strip of reef along which 100 m of transect tape is laid along the reef edge along a designated tide-corrected depth contour (< 1 m variation), typically at 3 – 6 m, but ranging from 1 to 12 m. A system of detailed maps and tide times is utilized to ensure consistent transect placement from year to year, although the precise placement and alignment of the tape within the narrow strip may vary from one visit to the next. ‘Point intercept transect’ (PIT) sampling is conducted at 0.5 m intervals along 20 m segments of the tape, wherein a SCUBA diver or snorkeler records the benthic category touched by a plumb-bob lowered beneath the point (Hodgson et al 2006; Hill and Wilkinson 2004). Sampled tape segments are separated end to end by ~ 5 m, providing independent replicate transects within the site, and a standard sampling effort depending on the number of segments. In most RCA sites there are 4 segments (160 points), but at a minority, limited reef area and/or logistic constraints permit the use of only three (120 points). We calculated the percentage of points in each category to provide an estimate of its percentage cover, and the co-efficient of variability (CV) among segments of dominant benthos to provide a measure of the site’s heterogeneity with respect to that category.

Accuracy

The inherent limits to the accuracy (closeness of estimates to the real percentage cover) of the PIT protocol were simulated in Excel 2010. This accuracy simulation does not take into account variability among observers or line placements, the effects of which on precision were investigated in field trials described below. For the simulation, the conceptual reef comprised seven randomly distributed benthic categories with covers of 50%, 25%, 12.5%, 5%, 4%, 2.5% and 1% (total 100%) sampled along a 20 m transect line using 50, 100, 250, 500 and 1000 regularly spaced points. For ease of computation, the seven categories were deemed to occur in contiguous blocks arranged along the transect line from 0 to 20 m and the position of each point on the line was selected at random. Function RAND() was used to generate the proportional distance from the start of the transect line that each sampling point was to be placed along the line, and the benthic category at that point was recorded. This process was repeated 20 times for each sample size. Scores for each category were
then used to determine the standard deviation (α) among cover estimates at each sample size. The increase in accuracy with increasing sample size was illustrated in plots, with particular reference to the RCA sample sizes of 40 points (1 x 20 m transect), 120 points (3 x 20 m transects) and 160 points (4 x 20 m transects).

**Study sites**

Field studies of precision and power were undertaken at two GBR sites (Nelly Bay and Moore Reef (Fig. 1) in April 2007, and one SEQ site (Shag Rock) in February 2014. These sites were chosen to represent major elements of the benthic variability across the > 70 routinely monitored sites (Fig 1), with the premise that knowledge gained about sources of measurement variability and power will guide the interpretation of the other sites not considered here (see Loder et al. 2015 for a summary).

**Nelly Bay:** The Nelly Bay site at Magnetic Island (146.85°E, 19.17°S; tropical; high and homogeneous coral cover) is a shallow fringing reef ~12 km from Townville. The reef drops steeply to an adjacent sea floor of silty sand at a depth or around 10 – 12 m. Turbidity at the site is naturally high during prevailing SE winds due to resuspension of sediments (Larcombe et al. 1995), and dredging of a nearby shipping channel periodically elevates suspended solids. Routine monitoring at Nelly Bay commenced in 2002, and it was among many reefs which were damaged by a cyclone in 2008.

**Moore Reef:** Moore Reef (146.20°E, 16.87°S; tropical; moderate and patchy coral cover) is a mid-shelf platform reef > 30 km from the nearest land and the site of dive and snorkel tourism operations that visit daily from Cairns, some 50 km distant. Visitors usually experience high water clarity and diverse fish and invertebrate life among the corals. Routine monitoring commenced in 2004 and the reef was damaged by a cyclone in 2011. Like Nelly Bay, the Moore Reef study site is a well-defined narrow strip along the reef edge.

**Shag Rock:** Shag Rock (153.53°E, 27.41°S; sub-tropical; with patchy fleshy algae and coral on a rocky substrate) is a popular subtropical dive and snorkel site a few hundred metres from Point Lookout at the northern tip of North Stradbroke Island some 50 km from Brisbane port. The site attracts visitors on the basis of its fishes, invertebrates and charismatic megafauna (e.g. grey nurse.

**Precision**

We define precision as the 95% confidence interval of the means of independent PIT estimates of benthic cover. Three potential effects on precision were investigated: 1) ‘observer error’ (differences among observers in their estimates using the same line transect), 2) ‘deployment error’ (differences in estimates caused by minor differences in transect placement within the same site); 3) differences in the composition and patchiness of the benthic substrata between sites (herein also referred to as ‘homogeneity’ and ‘heterogeneity’).

The PIT precision study was implemented on SCUBA by groups of 5 - 10 trained volunteer observers at the three study areas. Transect start and end points were marked with pegs or concrete blocks to ensure all observers sampled the same line, thereby eliminating any unintended line placement effect. In the ‘observer error’ treatment, observers swam the same line consecutively, so although observations were made at the same 0.5 m intervals, there was usually some current- or swell-induced lateral displacement or swaying of the line that meant the point was not always precisely in the same place on the reef. This lateral movement of the transect line was a potential contributor to differences among observers in their cover estimates, in addition to potential differences in identification of the category. At Nelly Bay and Shag Rock, this procedure was repeated by observers along a single transect line. At Moore Reef only, in addition to ‘observer’ effects, line placement effects were explicitly investigated by repeatedly re-deploying and retrieving the transect line at two trials within the study area (for details of study design see Table 1). Percent cover estimates were calculated based on the number of points scored for each of the 10 RC categories.

Potential observer and/or deployment effects on cover estimates were investigated using non-parametric ANOVAs, as preliminary analysis using Levene’s test had established that data for several benthic categories did not meet requisites of parametric ANOVAs. For the observer variability trials at Nelly Bay and Shag Rock we used the Kruskal-Wallis one-way ANOVA on the ranked scores of each benthic category. For the observer and deployment trials (Moore Reef) we used Friedman’s two-way ANOVA on scores ranked 1) within observers across all deployments, and 2) within deployments across all observers. Both tests return a statistic $X^2$, which is distributed according to
Chi Square. For plotting, imprecision in the estimates (i.e. the variability among observers and/or deployments) was expressed as 95% confidence limits around the mean. Given the non-normality of the data, these confidence intervals were derived in Excel 2010 using bootstrap resampling (N = 500) from the cover estimates of 1) each observer across all their own deployments (all three sites), and 2) each deployment across all observers who surveyed it (Moore Reef only).

Power

This analysis investigated variation in the benthic cover data collected by volunteers using the PIT protocol in order to determine if there were variations in the power of detecting changes. ‘Power’ is defined as the probability of detecting an effect of a certain magnitude (in this case, change in absolute percent cover). Power was computed for the Nelly Bay site (4 x 20 m transect segments; 6 observers), the Moore Reef site (two trials, each with 4 x 20 m transect segments; Trial 1: 4 observers on 6 deployments; Trial 2: 6 observers on 3 deployments) and the Shag Rock site (3 x 20 m transect segments; 10 observers). Power was computed for individual observers based on the mean and standard deviation of their estimates (DSS 2015) and plotted against effect sizes of 5 to 50%. Alpha (probability of incorrectly rejecting the null hypothesis that no effect of the nominated effect size is present) was set at 5%.

Results

Accuracy

The simulation showed that at sample sizes of 120 or 160 points, estimates are more accurate for low percentage cover than high in absolute terms and less accurate in relative terms. Thus (Fig. 2a), the inherent uncertainty of a 160 point estimate for a true cover of 50% is 50 ± 6%, whereas for a true cover of 1% it is 1 ± 0.6%. As a proportion of the true cover, and taking into account these inaccuracies both above and below the true value (Fig. 2b), these numbers are equivalent to ~25% of a 50% bottom cover and >100% of a bottom cover of 1%. i.e. the true value could be anything from about half to double the estimate for very sparse bottom cover, whereas for high cover, estimate and true cover are relatively much closer.

Precision
The imprecision introduced by the use of multiple observers and/or deployments was small compared to real differences between the sites (Fig. 3). Differences in hard coral cover among the three reefs were significant at p ≤ 0.05 (68% vs. 42% vs. 3%); there was more soft coral at Moore Reef (10%) than the other two reefs (1%); and there was more fleshy algae at Shag Rock (24%) than Nelly Bay (2%). However, algal cover at Moore Reef (8%) was not significantly different from either of the other reefs due to its wide 95% CIs (10 – 16%), attributable to both observer bias (p = 0.001) and deployment bias (p = 0.020; Table 2).

The variability represented in Fig. 3 also reflects differences in the patchiness of benthic categories on the three reefs. For example, the 68% hard coral cover at Nelly Bay was less patchy (CV among segments 11%) than the 42% cover at Moore Reef (among-segment CV 41%). At Shag Rock, the 24% cover of fleshy algae was relatively homogeneous (among-segment CV 27%). At Nelly Bay, all six observers estimated the percent cover of all ten benthic categories within narrow bounds (95% CI ≤ ± 5%; Fig. 4a). For hard coral, soft coral and sand (benthic categories with “o” in Fig 4a) ‘observer’ was a statistically significant contributor (p ≤ 0.05) to the very small variability around the means (Table 2), but it is unlikely that any ecological significance can be attributed to this minor effect.

At the more heterogeneous GBR site (Moore Reef), there was much greater variability among all cover estimates (95% CI up to ± 12%; Fig. 4b), and there were statistically significant ‘observer’ and/or ‘deployment’ effects in a number of benthic categories (Table 2). Affected categories included the ecologically important categories of hard and soft corals and fleshy algae, as well as rock, rubble, sand and silt. Because this analysis is based on ranks, these statistics indicate 1) that some observers routinely scored higher or lower than others for hard coral, fleshy algae and sponge and 2) that minor differences in the placement of the transect line (deployment) at a heterogeneous site, combined with sideways tape movement due to local currents, did influence the estimated percent cover of hard and soft coral, fleshy algae, rock, rubble and sand. At the subtropical reef in SEQ (Shag Rock, Fig. 4c), there was major variability among observers in the estimation of the dominant benthic categories fleshy algae (24 ± 12%) and rock (42 ± 16%), but for neither of these was ‘observer’ a statistically significant treatment. By contrast, there were significant observer effects on the very small variability in estimates of hard coral (4 ± 1%) and somewhat greater variability in sand (12 ± 8%). Since none of these benthic categories is difficult to identify, we attribute these nominal ‘observer’ effects to
sideways movement of the tape caused by variations in the oceanic swell between consecutive observers, which affects many SEQ sites (c.f. a greater importance of local reef currents in the GBR sites).

**Power**

The power analysis (Fig. 5) indicated differences among the three sites in the effectiveness with which PIT could detect indicative changes in the percentage cover of the dominant living benthos. In the most homogeneous hard coral dominated community (Fig. 5a), the 4 x 20 m transect protocol had > 90% power to detect 5 years of hard coral recovery (+20%) or moderate levels of coral bleaching or disease (-12% to -14%), irrespective of the identity of the observer. Where hard coral cover was more patchy (Fig. 5b), power to detect change varied substantially among observers, but any observer could, nevertheless, detect typical impacts of storms, crown-of-thorns starfish (*Acanthaster planci*) outbreaks or multiple disturbances (-32% to -42%; Osborne et al, 2011), or 10 years of continuous recovery (+40%; Done et al. 2010) with power of >90%. For the dominant but patchily distributed fleshy algae at Shag Rock (Fig. 5c), the 3 x 20 m transect protocol was very powerful, with all observers having ≥ 90% power of detecting a change of < ± 10%

**Interpretation of time series data at the precision study sites**

RCA records of average cover of dominant benthic groups through time at the Nelly Bay, Moore Reef and Shag Rock monitoring sites are presented in Fig. 6, with error bars representing variability due to observer and deployment effects derived in the precision studies. For Nelly Bay hard coral, there was strong upward trend from 2003 to 2008 (Fig. 6a); the small variability among observers (95% CI) and small effect size detectable (10% change in coral cover detectable with 90% power; Fig. 5a) indicate that this six year trend was real and that regional bleaching in 2005 did not substantially reduce coral cover at this site. The fall in hard coral cover was initiated by Cyclone Hamish which crossed passed near this site in 2009, and it continued falling until 2013. At RCA’s Moore Reef study site, which is situated in a sheltered part of the reef, there was a major (~ 30%) 9 year net increase in hard coral from 2005 – 2013 (Fig. 6b), despite a decline in coral cover on more exposed parts of the reef due to Cyclone Yasi in 2011 (reported by AIMS 2016). The minor inter-annual rises and falls at the RCA Moore Reef site from 2007 – 2009 are spurious, being less than measurement variability by criteria of both 95% CI (around ± 10%) and statistical power (effect size...
of 30% with 90% power). The 20% fall from 2013 – 2014 is likewise spurious, due to the extremely low power demonstrated by some observers to detect a 20% change in such a heterogeneous reef (Fig. 5b). RCA’s hard coral estimates for 2015 (55%) and 2016 (60%) collected after the present analysis, showed a resumption of the pre-2013 trend, suggesting that the spurious fall may in part be due to the 2013 estimate being a little high due to chance deployment and/or observer bias in this heterogeneous site. At Shag Rock (Fig. 6c), the decline in hard coral cover from 20% in 2009 to 7% in 2013 was real, being much greater than the among observer variability (around 3%), and of an order detectable with high power by all observers (Fig. 5c). However the subsequent 2013 – 2015 increase does not constitute evidence of coral recovery at that time, being of similar order of magnitude to the among-observer error bar.

The PIT protocol was also effective for soft corals and fleshy algae. At Nelly Bay (Fig. 6a), the 2003 – 2005 fall and the 2008 – 2009 rise in fleshy algae were substantially larger than among observer variability (6%), and soft corals were shown to remain rare throughout. At Moore Reef (Fig. 6b), the fall in soft corals (2004 – 2007) appears real, whereas subsequent rises and falls in soft coral were comparable to the error bars, and hence not reliably established. Likewise, the temporal variability in fleshy algae is too small to be reliably established. At Shag Rock (Fig. 6c), the net decline of soft coral cover from 2009 – 2015 is greater than observer variability and thus appears to be real. The increase of fleshy algae cover starting in 2009 and the fall following 2012 are also strongly demonstrated but the apparent 15 - 20% fluctuation from 2010 to 2012 is smaller than the > 20% variability among observers. However there was not a significant ‘observer’ effect on the estimation of fleshy algae at Shag Rock (Table 2), and the 3 transect protocol is very powerful (90% power in detecting this order of change). Given the ephemeral nature of fleshy algae, the major fluctuation is more plausible than it would be in a time series for hard coral cover.

**Discussion**

These analyses show that RCA’s application of the PIT protocol is highly effective, despite a) the potential for errors caused by inherent limits to accuracy of using only 120 – 160 point PITs, b) any patchiness in the distribution of reef benthos, c) minor inconsistencies in line placement, and d) the use of different observers. Potential sampling error was greatest where benthic patchiness and consequent susceptibility to observer- and placement-derived variability was greatest. In very patchy
sites, therefore, the PIT is sufficiently powerful to detect only moderate to large changes between successive years. In principle, using the same observer to lay the tape and do the assessment from one year to the next could decrease observer and deployment errors, but this practice would not align with the goals of the capacity-building program to be as broadly inclusive as practicable. Moreover, given that RCA training and volunteer support have continued to evolve since the original precision study was conducted in 2007, this study in essence establishes a worst-case scenario for measurement errors. With these caveats, it is nevertheless clear that this analysis has provided concrete guidance as to when changes in benthic cover over one to several years can safely be considered as a representation of real change, or when they should simply be acknowledged as sampling noise.

Collectively, Nelly Bay, Moore Reef and Shag Rock represent a good sample of the benthic variability across all RCA sites, and moreover, in practice ‘observer’ and ‘deployment’ are not as random as the present analysis assumes. In routine surveys from year to year, more highly trained team leaders who are often familiar with the site lay the 100 m tape, tending to minimise deployment error. Under these circumstances, the conclusions to be drawn from this study appear sound: the 120 – 160 PIT protocol can document major coral losses such as can be caused by major coral bleaching, storms or coral predators, and they can detect multi-year trends in benthic cover in either direction; they cannot robustly demonstrate subtle changes (such as annual net increase in coral cover at a healthy site, or minor losses due to bleaching or disease); the protocol is most effective in relatively homogeneous study sites where there is high power to detect change; in heterogeneous sites, where statistical power is lower and more variable among observers, single year changes in mean cover should be treated with a degree of circumspection. Inasmuch as the RCA data base allows routine inspection of variability among the replicate 20 m segments, it does provide the means of differentiating noise from signal over years to decades.

The basic PIT analysed here is not intended to document local impacts, such as fin or anchor damage to coral: the high inherent relative discrepancies at low bottom cover (Fig. 2b) would preclude the PIT from reliably detecting, for example, an ecologically or aesthetically important increase in fin damage to coral of a few percent. Recognising this limitation of PIT, RCA undertakes separate reef impact surveys of coral damage, diseases, predator scars and other localized impacts (Hill and Wilkinson 2004) to flag early warning signs of health issues such as coral disease or crown of thorns starfish.
While it was practical for RCA to test the precision (variability among estimates *per se*) of benthic cover estimates in the field, it was not practical to undertake the more demanding test of accuracy (closeness to the real mean). However the computer simulation did show that 120 – 160 points are inherently accurate (>91%), almost identical to a different simulation result reported in Marsh et al. (1994). There would be little to be gained by increasing the number of points beyond the current 160 points. Other accuracy simulations have shown that the absolute disparity of 20 m linear point intercept estimates from the true value was only 1.5% at 40% bottom cover, and < 1 % at coral cover of < 10% (Nadon and Stirling 2005). Those studies also found that, for coral cover of 5 – 25%, precision levelled off at around 80 – 100 points, well below the 120 - 160 points used by RCA.

There is an important difference between our field study of observer variability and ones based on PIT of digital images taken along a transect, and where different observers can record on exactly same point on the image. In such studies (e.g. Ninio et al. 2003) ‘observer error’ is entirely attributable to identification inconsistencies among observers. In our field-based precision trials, by contrast, ‘observer error’ is strictly speaking an ‘observer + line sway’ error, where the swaying of the deployed transect line is caused by passing swells or differences in localised tidal currents. The high training pass marks set by RCA minimise identification inconsistencies, and it is likely that ‘sway’ error much is a greater contributor than ‘observer’ mis-identification of categories.

With this study we have demonstrated that the RCA program delivers two vital aspects of public participation in research (Dickinson et al. 2010): the opportunity for the meaningful engagement of community volunteers and the collection of data of known quality. A broadening and deepening of understanding of coral reef issues across the broader community is manifest in the activities volunteers are trained to undertake. How far such awareness transmits from the citizen scientist volunteers into the broader community could in itself be a subject for follow-up research. In relation to data quality, our assessments of accuracy, precision and power suggest that these benthic data, along with allied data from reef impact monitoring (Loder et al. 2015) are useful in providing reef science and management stakeholders (including tourism, research, reef management, concerned citizens and community organisations) with indicators of ecological conditions at a spatial scale appropriate for site management and condition reporting. RCA’s use of repeated measures within the same circumscribed reef area from year to year provides greater confidence in measuring temporal change than would be achieved through the same effort applied to random sampling of a
less well-defined area. The 100 m length of the RCA study sites is congruent with the sorts of
distances provided by tourism operators for recreational divers on a single dive, and indeed RCA’s
overall monitoring data set provides a better index of reef states likely to be seen by recreational
divers than do the broader range of habitats monitored by government (Done et al. 2010; Osborne et
al. 2011; De’ath 2012). RCA is continuing to provide a unique set of time series, many of which now
extend for over a decade, for a significant and highly relevant sample of Queensland’s reefs and
shores. RCA’s benthic monitoring data thus represents an irreplaceable baseline against which
warming seas, other pressures and the effectiveness of management may be measured and used
undertake informed action with the community.

Acknowledgements

RCA is a registered charity, all of whose activities including the work described here are supported
by grants, public donations and in-kind services. However the research reported here did not receive
any specific grant from funding agencies in the public, commercial or not-for-profit sectors. We
thank the RCA Industry Champions for supporting access to long-term monitoring sites and the
hundreds of Reef Check Australia volunteers who supported field and office activities. Thank you to
Great Barrier Reef Marine Park Authority, Australian Marine Park Tourism Operators, Australian
Institute of Marine Science for technical advice and support. Special thanks to Craig Syms for
assistance with the design and analysis of the precision study, to Hugh Sweatman (Australian
Institute of Marine Science) and Richard Pearson (James Cook University) for commenting on drafts
of this manuscript, and to the certified RCA volunteers who conducted the trials reported here under
the supervision of CR, AH, and LF. The Shag Rock survey was made possible through the Point
Lookout Ecological Assessment Project conducted by University of Queensland Underwater Club
(UniDive) volunteers.

References

rpdetail.jsp?fullReefID=16071S&sampleType=MANTA


Figure Captions

Figure 1. RCA monitoring sites Queensland. Location detail and years of monitoring at Moore Reef, Nelly Bay and Shag Rock, sites of the precision study reported here are indicated. Black dots indicate location of reefs with one or more RCA monitoring sites along the Queensland coast. Standard study site used in routine monitoring is a narrow band within which a 100 m tape measure is haphazardly deployed. For Shag Rock, a non-standard length of 75 m was used due to space and logistic constraints. Within each 25 m section of the site, a 20 m section of the 100 m tape measure is sampled at 0.5 m intervals of each deployment.

Figure 2. Inherent limits of Point Intercept Transects (PIT) based on simulation study using increasing numbers of points. a) Absolute accuracy, where converging lines indicate the true cover as indicated ± standard deviation (α) of 20 PIT estimates of cover. b) Relative discrepancy = α/x%, where x is the real cover as indicated. In both panels, broken vertical lines indicate the number of points sampled by 1, 3 and 4 20 m transect segments. Graphs were smoothed and intermediate lines were omitted for clarity.

Figure 3. Means and 95% confidence limits of estimates of major living benthos at three study sites. Error bars represent variability (95% confidence limits) across multiple observers and deployments. Non-overlap of error bars indicates that the estimates are significantly different at p = 0.05. Nelly Bay (6 observers, 1 deployment); Moore Reef Trial 1 (4 observers, 6 deployments) and Trial 2 (7 observers, 3 deployments); Shag Rock (10 observers, 1 deployment).

Figure 4. Cover estimates (means) of Reef Check’s ten basic benthic categories. Error bars represent 95% confidence limits, derived in Excel 2010 using bootstrap resampling (N = 500) from estimates across all observers and all deployments. a) Nelly Bay (6 observers, 1 deployment) b) Moore Reef, Trial 1 (4 observers, 6 deployments) and Trial 2 (7 observers, 3 deployments) c) Shag Rock (10 observers, 1 deployment). Letters ‘o’ (observer) and ‘d’ (deployment) indicate significant treatment effects (from Table 2).

Figure 5. Statistical power of 20 m point intercept transects. Coloured lines indicate power of individual observers in detecting the indicated effect size, with alpha error set at 5%.
Nelly Bay: 4 transects and 6 observers. Moore Reef: 4 transects and 5-6 observers. Shag Rock: 3 transects and 10 observers. Power calculator used was

Insets show relative cover and heterogeneity of dominant living benthos at Nelly Bay (homogeneous), Moore Reef (heterogeneous) and Shag Rock (heterogeneous). Variability (range 0 – 100%) has been smoothed using a 3 point running mean of the ones and zeros scored by one observer for the indicated category at 0.5 m intervals (check marks on x axis) along 4 (Nelly Bay; Moore Reef) or 3 (Shag Rock) 20 m transects commencing at 25 m intervals along the 100 m transect line). CV refers to the co-efficient of variation among the indicated number of transects.

Figure 6. Time series of dominant benthic categories at the three study sites. Estimated percentage cover was based on results of one observer using one set of 4 x 20 m transects at Nelly Bay and Moore Reef, and one set of 3 x 20 m transects at Shag Rock. Error bars (95% confidence interval) represent the variability among observers (Nelly Bay and Shag Rock) and variability among observers and deployments (Moore Reef) established in the precision studies of 2007 (Nelly Bay; Moore Reef) and 2014 (Shag Rock). Timing of regional coral bleaching (2005), Cyclone Hamish (2009) and Cyclone Yasi (2011) (all GBR only) are indicated. * indicates lack of signal in RCA site, despite reported regional impacts.
### Table 1. Summary of sites used for precision studies and routine monitoring

<table>
<thead>
<tr>
<th></th>
<th>Nelly Bay</th>
<th>Moore Reef</th>
<th>Shag Rock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dominant living benthos at time of precision study</td>
<td>Homogeneous hard coral</td>
<td>Patchy hard coral</td>
<td>Patchy fleshy algae</td>
</tr>
<tr>
<td>Routine surveys (Fig. 6)</td>
<td>Since 2002</td>
<td>Since 2002</td>
<td>Since 2009</td>
</tr>
<tr>
<td>Number of 20 m segments per transect line (routine surveys)</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Depth of transect</td>
<td>5 m</td>
<td>5 m</td>
<td>6 m</td>
</tr>
<tr>
<td>Precision study date</td>
<td>April 2007</td>
<td>April 2007</td>
<td>Feb 2014</td>
</tr>
<tr>
<td>Treatment</td>
<td>Observer</td>
<td>Observer, Deployment</td>
<td>Observer</td>
</tr>
<tr>
<td>Precision study design (Fig. 4)</td>
<td>6 observers on 1 deployment</td>
<td>Trial 1: 4 observers on 6 deployments</td>
<td>Trial 2: 7 observers on 3 deployments</td>
</tr>
<tr>
<td>No of 20 m transects per deployment (precision study; Fig. 4)</td>
<td>6</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>No of 20 transects in power analysis (Fig. 5)</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 2: Observer and deployment effects on estimates of % cover of Reef Check Basic benthic categories, derived from non-parametric analyses of variance (ranked data). Null hypotheses Ho are that neither observer nor deployment has statistically significant effect on estimate of percentage cover. Results are from one trial at Nelly Bay (one transect deployment surveyed by five observers), one trial at Shag Rock (one transect deployment surveyed by ten observers) and two trials at Moore Reef (three transect deployments each surveyed by seven observers; six transect deployments, each surveyed by four observers).

p = probabilities under Ho that $X^2 \geq \text{Chi Square}$. Values of p ≤ 0.05 are indicated in bold text.

<table>
<thead>
<tr>
<th>Benthic category</th>
<th>Nelly Bay</th>
<th>Shag Rock</th>
<th>Moore Reef Trial 1</th>
<th>Moore Reef Trial 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observers (6) DF = 5</td>
<td>Observers (10) DF = 9</td>
<td>Observers (4) Deployments (6) DF = 3</td>
<td>Observers (7) Deployments (3) DF = 6</td>
</tr>
</tbody>
</table>

**Living substrata**

<table>
<thead>
<tr>
<th>Benthic category</th>
<th>Kruskal- Wallis test</th>
<th>Friedman’s test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard coral</td>
<td>$X^2$ 15.967</td>
<td>$X^2$ 19.054</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.010</td>
<td>$p$ 0.050</td>
</tr>
<tr>
<td>Soft coral</td>
<td>$X^2$ 15.117</td>
<td>$X^2$ 4.119</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.010</td>
<td>$p$ 0.950</td>
</tr>
<tr>
<td>Recently killed coral</td>
<td>$X^2$ 0.434</td>
<td>$X^2$ 3.123</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.980</td>
<td>$p$ 0.980</td>
</tr>
<tr>
<td>Fleshy algae</td>
<td>$X^2$ 1.925</td>
<td>$X^2$ 10.253</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.800</td>
<td>$p$ 0.500</td>
</tr>
<tr>
<td>Sponge</td>
<td>$X^2$ 5.948</td>
<td>$X^2$ 1.036</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.300</td>
<td>$p$ 0.990</td>
</tr>
<tr>
<td>Other</td>
<td>$X^2$ 1.257</td>
<td>$X^2$ 5.322</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.900</td>
<td>$p$ 0.900</td>
</tr>
</tbody>
</table>

**Non-living substrata**

<table>
<thead>
<tr>
<th>Benthic category</th>
<th>Kruskal- Wallis test</th>
<th>Friedman’s test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock</td>
<td>$X^2$ 6.030</td>
<td>$X^2$ 7.187</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.200</td>
<td>$p$ 0.700</td>
</tr>
<tr>
<td>Rubble</td>
<td>$X^2$ 8.782</td>
<td>$X^2$ 5.549</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.100</td>
<td>$p$ 0.800</td>
</tr>
<tr>
<td>Sand</td>
<td>$X^2$ 17.860</td>
<td>$X^2$ 18.533</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.010</td>
<td>$p$ 0.050</td>
</tr>
<tr>
<td>Silt</td>
<td>$X^2$ 4.952</td>
<td>$X^2$ 4.826</td>
</tr>
<tr>
<td></td>
<td>$p$ 0.900</td>
<td>$p$ 0.900</td>
</tr>
<tr>
<td>p</td>
<td>0.300</td>
<td>0.900</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>506</td>
<td>507</td>
<td></td>
</tr>
</tbody>
</table>
508
Figure 1

Queensland


Shag Rock (2009 – 2015)
Figure 2

(a) Absolute accuracy

(b) Relative discrepancy
Figure 3

Cover (%)

- Nelly Bay
- Moore Reef
- Shag Rock

Hard coral
Soft coral
Fleshy algae
Figure 5

(a) Nelly Bay Hard coral
4 x 20 m transects: CV = 0.11

(b) Moore Reef Hard coral
4 x 20 m transects: CV = 0.41

(c) Shag Rock Fleshy algae
3 x 20 m transects: CV = 0.27
Figure 6

(a) Nelly Bay

(b) Moore Reef

(c) Shag Rock

Legend:
- Hard coral
- Fleshy algae
- Soft coral

Events:
- Cyclone Hamish
- Cyclone Yasi
- Bleaching

Cover (%)