Title
Evaluating ecological states of rocky intertidal communities: A Best Professional Judgment exercise

Authors
Murray, SN
Weisberg, SB
Raimondi, PT
et al.

Publication Date
2016

DOI
10.1016/j.ecolind.2015.08.017

Peer reviewed

Evaluating Ecological States of West Coast Rocky Intertidal Communities: A Best Professional Judgment Exercise

*S.N. Murray[1], S.B. Weisberg[2], P.T. Raimondi[3], R.F. Ambrose[4], C.A. Bell[3], C.A. Blanchette[5], J.L. Burnaford[1], M.N. Dethier[6], J.M. Engle[5] M.S. Foster[7], C.M. Miner[3], K.J. Nielsen[8], J.S. Pearse[3], D.V. Richards[9], and J.R. Smith[10]

*Corresponding Author

[1]Department of Biological Science, California State University, Fullerton, CA 92834-6850 USA; [2]Southern California Coastal Water Research Project Authority, 3535 Harbor Blvd. #110, Costa Mesa, CA 92626 USA; [3]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Long Marine Lab, Santa Cruz, CA 95060 USA; [4]Department of Environmental Health Sciences, Box 951772, University of California, Los Angeles, CA 90095-1772 USA; [5]Marine Science Institute, University of California, Santa Barbara, CA 93106-6150 USA; [6]Friday Harbor Laboratories and Biology Department, University of Washington, 620 University Road, Friday Harbor, WA 98250 USA; [7]Moss Landing Marine Laboratories, 8272 Moss Landing Road, Moss Landing CA 95039 USA; [8]Romberg Tiburon Center for Environmental Studies, San Francisco State University, Tiburon, CA 94920 USA; [9]Channel Islands National Park, 1901 Spinnaker Drive, Ventura, CA 93001 USA;[10]Biological Sciences Department, California State Polytechnic University, Pomona, 3801 West Temple Avenue, Pomona CA, 91768 USA.

Key Words

Ecological Evaluation; Infauna; Rocky Intertidal; Ecological Communities; Best Professional Judgment; Anthropogenic Disturbance

Highlight Bullet Points

Judgments of the ecological states of rocky intertidal communities were performed

Agreement was less for rocky intertidal compared with polyhaline infaunal communities

Spatial heterogeneity is greater in rocky intertidal vs soft bottom benthic habitats

State is hard to judge from one-off data in the temporally-variable rocky intertidal

Macrophytes and sessile macroinvertebrates were emphasized in rocky shore evaluations

Abstract

A Best Professional Judgment (BPJ) exercise was performed on west coast rocky intertidal communities to: 1) determine the level of expert agreement achieved in evaluating the states of rocky intertidal communities, 2) compare the level of agreement with similar expert assessments for soft bottom infaunal communities, and 3) identify the data experts found most useful in making their judgments. Species-abundance and environmental data were provided to 14 experts who independently ranked communities from best to worst and assigned them to one of five categories based on the degree of deviation from an expected, natural biological state. Experts averaged 70.0% and 75.4% Euclidean Similarity in evaluating 11 and 12 communities from central and southern California and achieved correlations in best-to-worst rankings of 0.30 and 0.49 respectively. Less agreement was achieved compared with infaunal experts for two groups of west coast polyhaline communities (82.1% and 80.0%; *r* = 0.91 and 0.91), but agreement was better than for tidal freshwater (63.9%; 0.29) and mesohaline (65.0%; 0.38) infaunal assemblages. State evaluations of rocky intertidal communities present challenges that go beyond those for soft-bottom, infaunal communities. These include: 1) greater spatial heterogeneity in rocky intertidal habitats; 2) less ability of rocky shore experts to assign deviations from an expected natural state without direct knowledge of anthropogenic stressors; 3) differences in the characteristics of infauna compared with rocky intertidal macro-organisms; and 4) the sampling protocols employed to generate infaunal and rocky intertidal community data. Rocky intertidal habitats are subjected to multiple, natural community-altering disturbances that vary over space and time causing experts to request more detailed, physical habitat descriptions before independently characterizing the natural disturbance regime and developing an expectation for community composition. Because analyzed data were for a single sampling

event, the relationship between the examined community state and a temporal range of expected natural states had to be projected before final evaluation. Most rocky intertidal experts emphasized macrophytes and sessile macroinvertebrates in making state evaluations. This study underscores the difficulties in distinguishing deviations from an expected natural state on rocky shores working from “one-off” species abundance data and physical site descriptions without information on the sources and magnitudes of anthropogenic perturbation. Such difficulties are likely in other habitats subjected to multiple stressors, and high spatial and temporal variation, particularly where biological responses to natural and anthropogenic stressors are often similar and difficult to distinguish.

1.0. Introduction

Coastal managers often rely on species abundance data to evaluate the ecological states of biological communities and to interpret the extent of anthropogenic impacts. Although multivariate approaches, such as non-metric multidimensional scaling (Clarke and Gorley 2006), are powerful tools for differentiating community states, analyses based on biological data can be difficult to interpret, particularly when the effects of multiple potential stressors need to be considered in the context of large natural biological variation. Moreover, coastal mangers rarely have access to temporal data sets with the history needed to evaluate community state in the context of natural community dynamics.

Biotic indices that translate complex ecological data into simpler metrics are sometimes used as communication tools for representing community states. Such indices have gained acceptance from coastal managers for characterizing the degree of anthropogenic perturbation in soft-bottom infaunal communities (Weisberg et al., 1997; Díaz et al., 2004; Borja et al., 2009). In response to a call from The European Water Framework Directive (EC, 2000), efforts have been made to develop similar indices for phytoplankton (Revilla et al., 2009) and macroalgae (Wilkinson et al., 2007; Selig et al., 2007; Sfriso et al., 2009), including macroalgal indices for rocky shores (Bard 1998, Pinedo et al., 2007; Juanes et al., 2008). To be effective, biotic indices must be based on well understood responses of populations and communities to natural and anthropogenic impacts. However, a consensus on which population and community responses best serve as consistent and reliable indicators of natural and most types of anthropogenic stress or the utility of a biotic index for rocky shores is lacking (Murray et al., 2006).

By their nature, rocky intertidal communities offer several challenges to evaluations of community state. First, these communities occupy heterogeneous habitats, generally with considerable spatial and temporal variation in key abiotic environmental drivers. This leads to multiple possible community states, even for habitat patches within the same physical site. Second, rocky shores can be simultaneously subjected to natural physical (e.g., wave action, sand scour, substratum instability) and biological (e.g., predation) disturbances, which can be difficult to differentiate from anthropogenic (e.g., poor water quality, trampling, harvesting) perturbations. Third, the rocky intertidal zone is strongly influenced by tides, whose submersion and emersion regimes limit the shore positions that can be occupied by most species. This produces well known, vertical patterns of species abundances within a site and makes it essential that site comparisons are based on community samples taken over equivalent shore positions. Fourth, the orientation (slope, aspect) and relief (rugosity) of the rocks themselves have a strong influence on species distributions and abundances, within and among sites (Schoch and Dethier 1996).

A step towards advancing our understanding of anthropogenically-impacted communities is to determine the level of consensus achieved by experts when asked to identify a community's ecological state based upon the response signatures captured by biological data. Best Professional Judgment (BPJ) exercises can be used to make such evaluations and have been successfully employed to make state judgments in many fields (Meyer and Booker, 2001). In aquatic environments, BPJ has been used to evaluate index performance (Ranasinghe, et al. 2008) and to determine consistency among experts in judging a community's ecological state (Bay et al., 2007). BPJ exercises have been used successfully to assess expert consensus in identifying the ecological states of west coast polyhaline (18 to 30 ‰), mesohaline (5 to 18 ‰), and tidal freshwater (< 5 ‰) infaunal communities (Weisberg et al., 2008; Texeira et al., 2010; Thompson et al., 2012) but have yet to be used for these purposes on rocky shores.

Here we convened a team of U.S. Pacific coast experts to conduct the first BPJ exercise performed on rocky shore communities to determine: 1) the level of agreement among west coast experts in identifying the states of rocky intertidal communities; 2) how well this level of agreement compares with BPJ assessments of the ecological states of west coast infaunal communities (Weisberg et al., 2008; Thompson et al., 2012); and, 3) which biological attributes experts found most useful in making evaluations. Experts evaluated the states of rocky intertidal macro-organism communities from both the central and southern California coast, separately, using biological data commonly collected in rocky intertidal sampling programs: site-scale data representing the abundances of macrophyte (macroalgae and surfgrasses) and macroinvertebrate (invertebrates discernible in the field with the unaided eye) populations.

2.0. Methods

We used similar procedures for central and southern California rocky intertidal communities as those employed in BPJ exercises for west coast infaunal communities (Weisberg et al., 2008; Thompson et al., 2012). Individuals with expertise sampling U.S. Pacific coast rocky shores were invited to participate. Identical biological and physical environmental data sets for a range of sites were collected, standardized and given to each expert, and instructions for ranking and scoring site communities were provided so experts could prepare their evaluations.

2.1. Experts. Fourteen experts with 10 to 40 years of experience working on community-level, field sampling of U.S. Pacific coast rocky intertidal communities participated. Almost all had extensive field sampling experience in central California whereas only a few had similar sampling experience in southern California; one expert had the majority of experience outside of California, although at locations biogeographically similar to central California. Most experts routinely work with rocky intertidal population and community data and examine natural or anthropogenic drivers of community state. All 14 experts participated in the central California exercise while 13 experts submitted evaluations for southern California communities.

2.2. Sites. Experts were provided community data for 12 central California and 11 southern California sites (Table 1). The central California sites were distributed from Pigeon Point (37.19° N; 122.40° W) to Stairs (34.73° N; 120.62° W). Southern California sites were located south of Point Conception, a major biogeographic boundary, and were distributed from Old Stairs (34.07°N; 119.00°W) to Cabrillo Zone 1 (32.67°W; 117.25°W). All sites were located on the mainland and data were collected between 2000 and 2007.

Table 1. Names and Locations of Rocky Intertidal Sites Containing Macro-organism Communities Evaluated by Experts.

Site Number	Site Name	Latitude (°N) (DD.DD)	Longitude (°W) (DD.DD)
Central California Sites			
4	Pigeon Point	37.19	122.40
5	Año Nuevo	37.11	122.33
9	Hopkins Marine Station	36.62	121.91
11	Point Lobos	36.51	121.94
17	Hazard Canyon	35.29	120.88
18	Shell Beach	35.17	120.70
19	Stairs	34.73	120.62
22	Partington Cove	36.17	121.70
23	Lucia	36.01	121.54
24	Duck Pond	35.86	121.42
29	Terrace Point	36.95	122.06
30	Point Sierra Nevada	35.73	121.33
Southern California Sites			
1	Buck Gully	33.59	117.87
2	Cabrillo: Zone 1	32.67	117.25
3	Crystal Cove	33.57	117.84
4	Dana Point	33.46	117.71
5	Heisler Park, Laguna Beach	33.54	117.79
6	La Jolla Caves	32.85	117.27
7	Lechuza Point	34.03	118.86
8	Old Stairs	34.07	119.00
9	Paradise Cove	34.01	118.79
10	Scripps	32.87	117.25
11	Sequit Point	34.03	118.86

2.3. Data Sets. Data were obtained using common rocky intertidal community sampling procedures. At each site, a 30 m transect parallel to the tideline above the high tide zone was established to provide a baseline for locating 11 additional transects running perpendicular to the ocean; these 11 transect lines, separated by 3 m intervals, followed the contours of the shoreline. Cover data for macrophytes and macroinvertebrates were collected along the 11 transects using a point intercept method; intervals between points were adjusted along transects with respect to the topographic features and extent of the sampled habitat in order to achieve 100 points per transect.

In addition to point data, counts were obtained for mobile macroinvertebrates and converted to density. This was accomplished by counting the mobile macroinvertebrates contained in three 50 cm x 50 cm quadrats randomly placed in high, mid, and low zones along each of the 11 transects (n = 33 quadrats). Estimates of the tidal heights of sampled points and quadrats were made and measures of selected physical environmental features performed. Biological data in this exercise

Table 2. Biological and Environmental Data Provided for Evaluations of Central and Southern California Rocky Intertidal Communities. Qualitative Metrics (e.g., very low, low, moderate, high, very high) Were Used for Parameters Indicated with an Asterisk (*). Degree of Freshwater Influence Was Assessed by Providing Source and Proximity.

Biological Data	**Environmental Data**
Mean Biological Cover	*Site Location*
Mean Abiotic Cover (Bare Rock, Sand, Tar)	Site Name
	Latitude
Macrophytes	Longitude
Mean Cover for Site	Biogeographic Affinity
Mean Cover by 0.3m Tidal interval	
	Substratum
Macroinvertebrates	Substratum Geological Formation
Mean Cover for Site	*Substratum Character (Degree of Consolidation)
Mean Cover by 0.3m Tidal interval	*Susceptibility to Substratum Breakout
Mean Density for Site	*Substratum Relief
Mean Density by 0.3m Tidal interval	Substratum Slope (Degrees)
	Primary Substratum Type (e.g., Bedrock, Boulders)
Species and Taxon Diversity (Separately for Cover and Density Data)	
Number of Sampled Taxa	*Physical Disturbance Agents*
Total Cover or Number of Organisms	*Degree of Wave Exposure
Simpson's $1 - \lambda$ Index	Wave Exposure (Primary Direction)
Shannon's H' Index (ln)	Degree of Freshwater Influence
Pielou's J' Index (ln)	*Degree of Sand Influence
Margalef's D' Index	
	Other
Species and Taxon Characteristics	Protection Status
Classification	*Degree of Human Visitation
Functional Group	Site Overview Photographs
Trophic Group (Macroinvertebrates)	

consisted of species abundances (cover and density) for each site, presented in a species by site matrix in an Excel spreadsheet, plus several summary and diversity calculations made from these data; experts were also given key environmental features and the sampling date for each site (Table 2).

2.4. Exercises. Unlike the BPJ exercises for Pacific coast infaunal communities, which relied on a single evaluation (Weisberg et al., 2008; Thompson et al., 2012), three trials were performed. Adjustments were made in the type and amount of information provided after each of the first two trials to clarify instructions and address questions raised during post-trial discussions. Thus, experts were informed by discussions stemming from the previous trial before submitting their responses for the ensuing trial. Between trials, experts were given several weeks to individually process post-trial discussion and then were provided biological and environmental data sets and asked to submit their own independent evaluations for the next trial.

For Trial 1, experts were given data from 31 central California intertidal communities. The names and locations of sites were not provided prior to evaluations, requiring experts to make assessments using only site-specific biological and environmental data; no indication of the sources or magnitudes of anthropogenic stressors were included. Experts were asked to assign each site's state to one of five categories regardless of the nature of perceived disturbance: 1) undisturbed; 2) largely undisturbed; 3) neutral; 4) moderately disturbed; and 5) strongly disturbed, and to identify the five sites believed to be most strongly influenced by anthropogenic disturbance. Responses were summarized, presented to experts, and discussed during a post-exercise meeting.

For Trial 2, 12 of the original central California sites were selected for further analysis; biogeographically different communities from 11 southern California sites were added to the exercise to potentially enlarge the range of anthropogenically-impacted conditions and to increase the diversity of biological communities subject to evaluation. For the southern California sites, experts were not provided site locations, but for central California the site locations and expert rankings in the first trial were known from post-Trial 1 discussion. Experts were again asked to independently assign each site to one of five disturbance categories, without information on anthropogenic impacts, and to rank the sites from least to most disturbed (separately for central and southern California sites). No attempts were made to distinguish anthropogenic from other forms of disturbance in these evaluations because of difficulties in making this distinction revealed during post-Trial 1 discussion.

The same set of sites was used in Trial 3, but besides the biological data, experts were provided additional environmental data including photographs and verbal descriptions of the physical characteristics of each site. This was done in response to feedback following Trial 2 where the experts felt the need for more information on site characteristics to better allow them to establish an expected community state. Post-Trial 2 discussions also revealed problems scoring communities exposed to different forms and magnitudes of natural disturbance (e.g., wave exposure) based on a “one-off” biological data set. Thus, for Trial 3, the disturbance scale was modified to reflect the degree to which the biological data match expectations because site biology is known to differ for sites with different environmental features and natural disturbance regimes. This approach required experts to first categorize sites based on their exposure to natural environmental conditions and then to use biological data to determine the degree to which the observed community state deviated from expectations for that type of site. Again, information on the sources and magnitudes of anthropogenic disturbance were not provided. For Trial 3, the scale was refined as follows: 1) undisturbed or within the envelope of states that characterize the "best that it could be" potential condition for a site of this type; 2) largely undisturbed or near to but outside a state that characterizes the "best that it could be" potential condition for a site of this type; 3) neutral; 4) moderately disturbed or removed from a state that characterizes the "best that it could be" potential condition for a site of this type; and 5) strongly disturbed or far removed from a state that characterizes the "best that it could be" potential condition for a site of this type. No attempt was made to adjust scores based on whether deviations from the expected state were due to anthropogenic factors or extreme natural disturbance events.

To determine the most informative elements of the biological data, experts identified and rated the usefulness of selected biological attributes in making their evaluations. Again, a five point scale was employed: 1) provides critical information of primary importance; 2) provides valuable information of importance; 3) provides information of value; 4) provides information but used as a secondary factor; and 5) provides little, if any, information and of limited or no use. Lastly, experts rated the usefulness of the non-biological variables made available for characterizing site types based on natural environmental features. This was done using the same five point scale.

2.5. Data Analyses. Descriptive statistical summaries [means ± 1 SD and CV (%)] were used to assess the aspects of the biological data most useful in making state evaluations and of the different environmental parameters used to characterize site types. Two approaches were used to determine the level of agreement in using the biological data to judge community states: expert rankings from "most" to "least" disturbed and assigned scores using the final five point disturbance scale. Separate analyses were performed for central California and southern California and final Trial 3 results were used to determine the level of agreement and for comparisons with BPJ exercises for polyhaline, mesohaline, and tidal freshwater infaunal communities (Weisberg et al., 2008; Thompson et al., 2012). Community state scores were compared across the three rocky intertidal trials to examine changes in the level of agreement as the exercise progressed

2.5.1. Rankings. PRIMER-e v6 (Clarke, 1993; Clarke and Gorley, 2006) was used to compute Spearman's correlation coefficients (*r*) between each rocky intertidal expert pair based on Trial 3 responses. Significance of *r* was determined using two-tailed probability tables and $\alpha \leq 0.05$ and ≤ 0.01. The degree of correlation among experts (mean, maximum, minimum *r*, number of pairs with significant and negative *r* values) was then compared with *r* values reported for west coast infaunal communities by Weisberg et al. (2008) and Thompson et al. (2012).

2.5.2. Deviation-from-Expectation Disturbance Scores. Two methods were used to analyze agreement among experts in scoring sites: 1) mean scores and CV for each community calculated among all experts and for the responses of each expert; and, 2) PRIMER-e v6 analyses based on pairwise similarities calculated between each expert pair followed by cluster and non-metric multidimensional scaling (MDS) analyses; these were performed treating the experts as samples and the sites as variables (Clarke, 1993; Clarke and Gorley, 2006). For the similarity analyses, a matrix of between-expert scores was constructed separately for central and southern California using the Euclidean Distance coefficient converted to a similarity percentage by the formula: Similarity (%) = $100\ (1\text{-}ED/4\ n^{0.5})$, where ED = Euclidean Distance and n = number of total expert pairs. Experts (samples) were then grouped using cluster analysis (group average) and subjected to MDS. Cluster groups (≥ 60% and 80%) were overlaid on MDS plots to show patterns of expert similarity. For comparison purposes, pairwise similarities were calculated and used in the same cluster and MDS analyses from data reported by Weisberg et al. (2008) and Thompson et al. (2012) for west coast infaunal communities.

Analyses of expert responses for the three trials were based only on the 12 central and 11 southern California sites assessed during Trial 3 and limited to comparisons of mean site disturbance scores and mean pairwise similarity values computed from these scores. Differences among the three central California and two southern California trials were tested statistically using repeated measures analyses with site and expert being categorical factors and trial being the repeated measure. Within subject effects were modeled using a Pillai Trace F approximation when appropriate (> 2 trials). Because of the lack of independence among pairwise similarity (%) values within each trial, only descriptive statistics (means ± 1SD) are presented to compare trial-to-trial differences in expert responses.

3.0. Results

3.1. Trial-by-Trial Results for Rocky Shores. Disturbance scores changed significantly for central (Multivariate Repeated Measures Analysis; $p < 0.001$) and southern (Univariate Repeated Measures Analysis; $p < 0.001$) California rocky intertidal communities following post-trial discussions. For central California, the mean site score did not differ significantly between the first two trials (3.09 vs 3.04; $p = 0.328$) but was significantly ($p < 0.001$) lower for Trial 3 (2.04), probably due to changes in and improved understanding of the disturbance scale; similar results occurred for southern California between the two trials (3.31 vs 2.94). For central California, mean similarity increased progressively across the three trials (63.5%; 70.7%; 75.4%) while variation decreased (CV = 9.4%; 8.4%; 6.2%), indicating improvements in the level of expert agreement. For southern California, a small increase in mean similarity also was observed (67.4% vs 70.0%) between the two trials along with a small increase in CV (7.0% vs 8.8%).

3.2. Central California Rocky Intertidal Communities – Trial 3. Expert agreement in ranking communities, as measured by Spearman's Correlation Coefficient *r*, averaged 0.49 for central California (Table 3). Thirty-two of 91 calculations resulted in an *r* value ≥ 0.60; 39.6% of *r* values were significant at $p \leq 0.05$ and 12.1% at $p \leq 0.01$. Only one expert showed negative correlations with peers and only in two cases; all other correlations were positive. State scores for central California across all experts ranged from 1.36 to 2.86 with an overall mean of 2.04 (Table 4). Five of the 12 communities were scored the same by a majority of experts but none received the same score from all experts. Only one community received at least one score of 5, whereas ten were given scores of 1 by at least one expert. The mean CV averaged 35.4% across all communities. The average state score for an individual expert ranged from 1.50 to 2.58 and the mean CV from 13.9% to 69.3% (Table 4).

Table 3. Central California. Spearman's Correlation Coefficients (*r*) Between Rocky Intertidal Experts Based on Site Community Rankings Using Trial 3 Results. Red Values - Significant Correlations ($p \leq 0.01$); Blue Values - Significant Correlations ($p \leq 0.05 \geq 0.01$); Yellow-shaded cells, negative correlations; Based on Two Tailed Tests.

	Experts (n=14)													
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**
A														
B	**0.806**													
C	**0.788**	**0.699**												
D	**0.921**	**0.755**	**0.762**											
E	0.536	0.497	**0.685**	0.552										
F	0.532	0.448	0.350	0.524	0.315									
G	0.280	0.021	0.014	0.175	-0.049	0.427								
H	**0.602**	0.252	0.259	0.497	0.105	0.462	**0.720**							
I	**0.592**	**0.622**	0.490	**0.685**	0.441	**0.713**	0.434	0.483						
J	0.504	0.091	0.308	0.280	0.182	0.406	**0.650**	**0.825**	0.175					
K	**0.602**	0.545	0.406	**0.650**	0.140	0.469	**0.629**	**0.601**	**0.650**	0.392				
L	**0.732**	0.545	0.503	**0.664**	**0.706**	0.517	0.448	**0.650**	**0.664**	0.490	0.427			
M	**0.655**	0.559	0.552	**0.818**	**0.601**	0.538	0.175	0.231	**0.664**	0.000	0.448	**0.594**		
N	**0.673**	**0.587**	**0.832**	**0.741**	**0.699**	0.503	-0.091	0.301	0.462	0.315	0.385	0.462	**0.587**	

Table 4. Central California. Rocky Intertidal Experts Disturbance Scores for Site Communities Using Trial 3 Results. Scores Are on a 1 to 5 Scale With 1 Being Least and 5 Being Most Disturbed.

	Experts (n = 14)															
Site	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**Mean**	**CV (%)**
4	2	2	3	2	1	2	2	2	2	2	1	1	1	2	1.79	32.4
5	2	2	2	3	2	2	2	2	2	3	2	2	2	2	2.14	16.9
9	1	1	2	2	1	2	1	2	1	1	1	1	1	2	1.36	36.6
11	2	1.5	3	3	2	2	2	2	3	3	2	2	1	2	2.18	27.9
17	1	1.5	2	3	2	2	1	2	3	1	1	2	1	2	1.75	40.0
18	2	3	3	3	3	3	2	2	3	2	5	2	4	3	2.86	30.3
19	2	2	3	3	1	3	2	2	3	2	4	1	1	2	2.21	40.3
22	1	1	2	1	1	3	2	2	1	3	1	2	1	2	1.64	45.3
23	1	1	2	2	1	1	3	2	2	2	2	2	1	3	1.79	39.2
24	3	2	4	4	2	2	3	3	2	2	3	3	1	3	2.64	31.9
29	3	3	2	3	1	4	3	2	3	2	3	3	2	2	2.57	29.4
30	1	1	1	2	1	1	2	2	1	1	2	1	4	2	1.57	54.2
Mean	1.75	1.75	2.42	2.58	1.50	2.25	2.08	2.08	2.17	2.00	2.25	1.83	1.67	2.25	2.04	35.4
CV (%)	43.1	41.3	32.8	30.7	44.9	38.5	32.1	13.9	38.5	36.9	57.2	39.1	69.3	20.1	38.5	

Table 5. Central California. Pairwise Similarities (Based on Conversion of Euclidean Distance Calculation) Between Rocky Intertidal Experts Based on Site Community Disturbance Scores for Trial 3. Red Values - Similarities ≥ 80 %; Blue Values - Similarities ≤ 60 %.

	Experts (n=14)													
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**
A														
B	**88.6**													
C	77.2	74.5												
D	75.0	73.5	**82.3**											
E	78.4	**81.6**	70.2	68.6										
F	77.2	**80.2**	75.0	73.0	68.6									
G	**82.3**	79.0	77.2	75.0	72.1	75.0								
H	**82.3**	**80.2**	**82.3**	79.6	78.4	77.2	**85.6**							
I	78.4	**81.6**	78.4	**80.9**	75.0	78.4	76.1	78.4						
J	78.4	76.6	76.1	70.2	75.0	76.1	**80.9**	**80.9**	75.0					
K	71.1	74.5	71.1	73.0	66.9	71.1	71.1	69.4	74.0	65.4				
L	**83.9**	**81.6**	74.0	74.0	79.6	76.1	**83.9**	**83.9**	77.2	79.6	66.9			
M	66.9	72.5	**57.3**	**59.8**	73.0	62.5	66.9	68.6	63.2	63.2	66.9	66.2		
N	77.2	77.8	**82.3**	79.6	76.1	75.0	**85.6**	**89.8**	78.4	78.4	73.0	**80.9**	68.6	

Euclidean similarity between experts averaged 75.4% and ranged from 57.3% to 89.8% for central California (Table 5). Similarity between experts was ≥ 80.0% in 21 and ≤ 60.0% in 2 of 91 cases; the mean CV was 8.3%. As depicted in MDS plots, cluster analysis revealed that all experts grouped together at ≥ 60.0% similarity; two groups of experts, one consisting of 2 and the other of 6, achieved ≥ 80.0% similarity in their state scores (Figure 1*a*).

3.3. Southern California Rocky Intertidal Communities – Trial 3. The mean *r* for southern California ($r = 0.30$) was lower than for central California and only 15 of 78 calculations resulted in an *r* value ≥ 0.60 (Table 6); 16 of the 78 correlations were significant at $p \leq 0.05$ and only 3 at $p \leq 0.01$. Negative *r* values were observed for three experts in a total of nine cases; all other correlations were positive. Southern California state scores were generally higher than central California and ranged from 2.38 to 4.08 with a mean of 2.94 (Table 7). Only one of the eleven communities was scored the same by a majority of experts; none received the same score from all experts. Unlike central California, at least one expert assigned a score of 5 to four

Table 6. Southern California. Spearman's Correlation Coefficients (*r*) Between Rocky Intertidal Experts Based on Site Community Rankings Using Trial 3 Results. Red Values - Significant Correlations ($p \leq 0.01$); Blue Values - Significant Correlations ($p \leq 0.05 \geq 0.01$); Yellow-shaded cells, negative correlations; Based on Two Tailed Tests.

	Experts (n=13)												
	A	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**
A													
C	0.282												
D	**0.682**	**0.645**											
E	0.309	0.255	0.409										
F	**0.645**	0.591	**0.736**	0.364									
G	0.146	0.296	0.579	0.155	0.305								
H	0.564	0.591	**0.709**	**0.718**	0.582	0.364							
I	-0.245	0.145	0.045	0.591	0.291	0.164	0.382						
J	0.091	0.236	0.464	0.082	**0.618**	0.187	0.036	0.400					
K	0.427	0.473	**0.627**	0.136	**0.609**	0.096	0.582	0.136	0.455				
L	**0.601**	0.264	**0.683**	**0.651**	**0.743**	0.429	**0.820**	0.465	0.342	**0.610**			
M	-0.009	-0.209	-0.073	0.327	0.036	0.159	0.118	0.518	0.127	0.100	0.287		
N	0.109	0.118	-0.227	0.318	-0.100	**-0.674**	0.209	0.109	-0.300	0.091	-0.087	0.209	

of the eleven communities and only five communities were given at least one state score of 1. The mean CV averaged 30.2%. The average score for an individual expert scoring all communities ranged from 2.18 to 3.55 and the mean CV from 9.8% to 56.8% (Table 7).

For southern California, Euclidean Similarity between experts was less than for central California, averaging 70.0% and ranging from 48.9% to 92.5% (Table 8). Similarity between expert pairs was ≥ 80.0% in 11 (14.1%) and ≤ 60.0% in 8 (10.3%) of 78 cases; the mean CV was 12.6%. As depicted in MDS plots, cluster analysis revealed that all experts grouped together at ≥ 60.0% similarity. However, only one group of four experts clustered at ≥ 80.0% (Figure 1*b*).

Table 7. Southern California. Rocky Intertidal Experts Disturbance Scores for Site Communities Using Trial 3 Results. Scores Are on a 1 to 5 Scale With 1 Being Least and 5 Being Most Disturbed.

	Experts (n = 13)														
Site	**A**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**Mean**	**CV (%)**
1	5	3	4	3	5	3	3	4	3	2	4	5	3	3.62	26.6
2	2	3	3	1	4	2	3	3.5	4	4	2	4	3	2.96	32.7
3	2	2	3	4	2	1	3	4	3	1	2	5	3	2.69	43.9
4	2	3	3	1	2	3	3	3	2	1	1	4	3	2.38	40.3
5	3.5	2	3	1	4	2	3	3	3	4	3	4	3	2.96	29.6
6	3	5	5	4	5	3	4	4	3	5	4	4	4	4.08	18.6
7	3	3	3	2	2	2	3	3	2	4	2	4	4	2.85	28.1
8	3	3	4	2	4	2	3	3.5	3	5	2	2	3	3.04	30.4
9	3	2	3	1	3	3	3	2	3	1	2	2	3	2.38	32.2
10	3	2	3	2	3	2	3	3.5	2	2	2	3	3	2.58	22.2
11	4	4	3	3	4	2	3	3	2	2	2	2	3	2.85	28.1
Mean	3.05	2.91	3.36	2.18	3.45	2.27	3.09	3.32	2.73	2.82	2.36	3.55	3.18	2.94	30.2
CV (%)	29.8	32.4	20.0	53.5	32.7	28.5	9.8	18.2	23.7	56.8	39.1	31.8	12.7	29.9	

Table 8. Southern California. Pairwise Similarities (Based on Conversion of Euclidean Distance Calculation) Between Rocky Intertidal Experts Based on Site Community Disturbance Scores for Trial 3. Red Values - Similarities ≥ 80 %; Blue Values - Similarities ≤ 60 %.

	Experts (n=13)												
	A	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**
A													
C	71.6												
D	75.9	**80.1**											
E	62.1	68.0	62.3										
F	75.9	71.8	**80.1**	**53.5**									
G	70.6	72.8	68.0	65.5	60.8								
H	77.1	**81.5**	**86.9**	66.3	73.9	75.0							
I	73.9	75.3	**85.4**	64.8	74.2	64.8	**85.4**						
J	70.6	71.8	75.0	64.6	69.8	75.0	**81.5**	76.5					
K	**56.5**	65.5	64.6	**50.6**	62.3	**56.0**	63.8	61.0	63.8				
L	74.7	71.8	70.8	73.9	68.0	77.4	73.9	69.2	76.2	62.3			
M	62.1	60.8	69.8	**52.9**	60.8	**54.8**	70.8	77.7	65.5	**48.9**	**59.4**		
N	75.9	**80.1**	**84.9**	63.8	70.8	71.8	**92.5**	**83.6**	77.4	64.6	70.8	71.8	

3.4. Comparisons of BPJ Results for Rocky Intertidal and Infaunal Communities.

BPJ studies with infaunal experts produced strong agreement in evaluations of polyhaline communities (Weisberg et al., 2008; Table 9). Mean *r* values calculated from rankings of southern California and San Francisco Bay polyhaline communities were 0.91 and 100% of the reported *r* values were significantly correlated at $p \leq 0.01$; there were no cases with negative correlations between expert rankings. This high level of agreement for polyhaline soft bottom communities was reinforced by disturbance scores. Majority agreement was achieved for 100% of the communities in both geographic regions (Weisberg et al., 2008). The mean CV among experts across all communities was 22.2% for San Francisco Bay and 20.1% for southern California. Our pairwise similarity calculations averaged 82.1% (CV = 9.6%) for San Francisco Bay and 80.0% (CV = 5.3%) for southern California; the maximum similarity achieved was 100% for both regions, whereas the minimum was 73.4% and 71.4% respectively. All but 3 experts for San Francisco Bay and 2 for southern California clustered together at ≥ 80% similarity (Figure 1*c*,*d*).

Although less than observed for polyhaline infaunal assemblages, most comparisons suggest that expert groups achieved greater consensus when evaluating central California rocky intertidal communities compared with mesohaline and tidal freshwater infaunal assemblages and southern California rocky intertidal communities (Table 9). The mean correlation between expert rankings was greatest ($r = 0.49$) and the percentage of negative correlations least (2.2%) for central California rocky intertidal communities than for the other non-polyhaline community types. Further, greater mean similarity (75.4%) between experts was observed for central California rocky intertidal communities and a higher percentage of experts clustered more closely together (Figure 1*a*). Evaluations of mesohaline and tidal freshwater infaunal (Thompson et al., 2012) assemblages and southern California rocky intertidal community states generally produced similar levels of agreement (Table 9). Greater mean r was found for rankings of mesohaline infaunal assemblages ($r = 0.38$) compared with southern California rocky intertidal communities ($r = 0.30$), and tidal freshwater infaunal ($r = 0.29$) assemblages. More negative correlations were observed between rankings of the mesohaline (17.9% of cases) and tidal freshwater communities (19.0%) compared with southern California rocky intertidal (11.5%) communities. Experts achieved majority agreement at a higher level when scoring tidal freshwater communities (85.0% of cases) compared with mesohaline (45.0%) and southern California rocky intertidal (9.0%) assemblages. However, our similarity analyses indicated greater agreement among experts in evaluating southern California rocky intertidal communities (mean similarity = 70.0%) compared with mesohaline (65.0%) and tidal freshwater (63.9%) infaunal assemblages (Figure 1*e*,*f*).

Table 9. Expert Evaluations for Central and Southern California Rocky Intertidal Communities, Polyhaline Infaunal Communities (Southern California and San Francisco Bay), Mesohaline Infaunal Communities (San Francisco Bay), and Tidal Freshwater Infaunal Communities (San Francisco Delta). Polyhaline Results Reported or Calculated from Weisberg et al. (2008) and Mesohaline and Tidal Freshwater Results From Thompson et al. (2012). Ranking Summaries Based on Comparisons of Spearman's Correlation Coefficient (*r*). Disturbance Score Summaries Based on Expert Scores and Euclidean Similarity Calculations.

Parameter	Polyhaline Infaunal San Francisco Bay	Polyhaline Infaunal Southern California	Mesohaline Infaunal San Francisco Bay	Tidal Freshwater Infaunal San Francisco Delta	Rocky Intertidal Macro-organisms Central California	Rocky Intertidal Macro-organisms Southern California
Number of Sites, Experts, and Possible Expert Comparisons						
Number of Sites Evaluated	11	24	20	20	12	11
Number of Experts	9	9	8	7	14	13
Total Expert Comparisons	36	36	28	21	91	78
Ranking Summaries						
Mean *r*	0.91	0.91	0.38	0.29	0.49	0.30
Maximum *r*	1.00.0	0.96	0.81	0.94	0.92	0.82
Minimum *r*	0.81	0.80	-0.27	-0.46	-0.09	-0.67
Pairs (%) Correlated at $p \leq 0.05$	100.0	100.0	57.1	57.1	39.6	20.5
Pairs (%) Correlated at $p \leq 0.01$	100.0	100.0	35.7	28.6	12.1	3.8
Pairs (%) With Negative *r*	0	0	17.9	19.0	2.2	11.5
Disturbance Score Summaries						
Sites (%) With 100 % Expert Agreement	18.2	4.2	0	0	0	0
Sites (%) With Majority Expert Agreement	100.0	100.0	45.0	85.0	41.7	9.0
Mean CV (%) Among Experts	22.2	20.1	30.2	31.6	35.4	30.2
Mean Similarity (%)	82.1	80.0	65.0	63.9	75.4	70.0
Similarity CV (%)	9.6	5.3	12.2	16.5	8.3	12.6
Maximum Similarity (%)	100.0	88.0	78.9	81.7	89.8	92.5
Minimum Similarity (%)	73.4	71.4	47.8	44.7	57.3	48.9

3.5. Useful Elements of Biological Data Sets. Most rocky intertidal experts relied on the abundances of macrophytes and sessile macroinvertebrates in making their assessments; abundances of mobile macroinvertebrates were generally found to be less useful (Tables 10, 11). Of lesser importance were overall biotic cover patterns, biological diversity and community level analyses, metrics and approaches commonly used to compare and analyze communities based on data sets for all populations. Experts identified disturbance indicators to be high abundances of small, fast-growing, opportunistic algae such as *Ulva* spp., small, red turf-forming and green filamentous algae and low-lying, crustose seaweeds. Macrophytes such as *Scytosiphon* spp. and *Petalonia* spp. and the anemone *Anthopleura elegantissima* were used to characterize sand-disturbed habitats. Experts focused on high abundances of upper shore rockweeds, lower shore kelp and other large brown seaweeds, and surfgrasses as low disturbance indicators. In addition, high abundances of selected, larger mobile invertebrates extracted from rocky shores by humans, such as black abalone and owl limpets, were considered as possible indicators of low site impacts.

Table 10. Importance of Biological Attributes Used by Rocky Intertidal Experts in Making Determinations of Community State. Lower Values Indicate Greater Importance. Reported are Means (± 1 SD) Based on the Following Scale: 1) Provides Critical Information of Primary Importance; 2) Provides Valuable Information of Importance;3) Provides Information of Value; 4) Provides Information But Used as a Secondary Factor; and 5) Provides Little if any Information; of Limited or No Use. Data Collected Separately for Central and Southern California Rocky Intertidal Assessments but Pooled Because of Consistency in Results.

Biological Attribute	Mean Response	± 1SD
Abundances of Species Groups or Selected Taxa	1.18	0.37
Overall Cover Patterns	1.71	0.64
Biological Diversity	2.79	0.99
Community Level Analyses	3.07	1.27
Species Distributions by Tidal Elevation	3.21	1.19

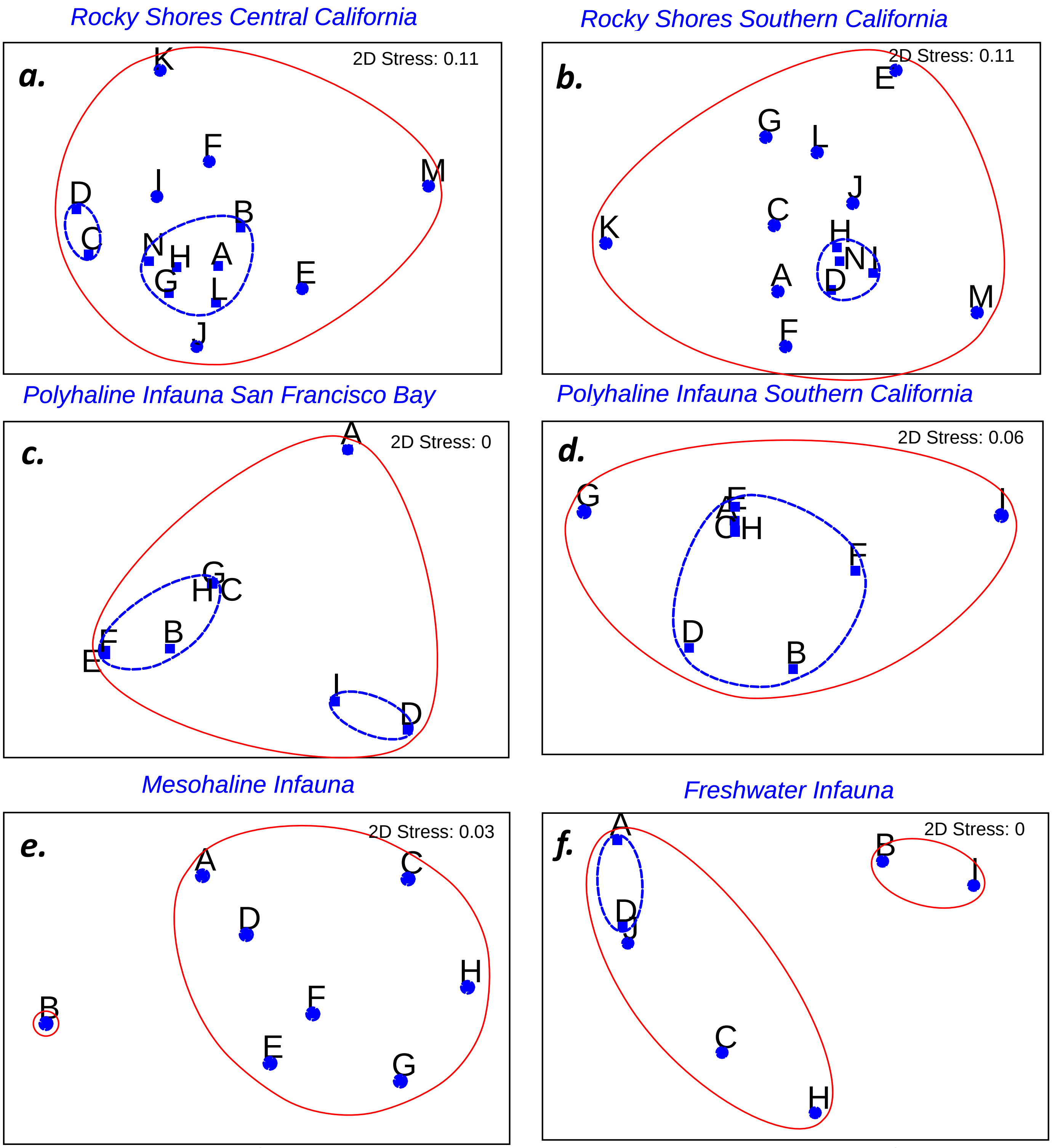

Figure 1. MDS plots of experts based on Euclidean Distance calculated as % similarity for Central California (*a*) and Southern California (*b*) Rocky Shores, Southern California (*c*) and San Francisco Bay (*d*) Polyhaline, San Francisco Bay Mesohaline (*e*) and San Francisco Delta Tidal Freshwater (*f*) Infaunal Communities. Polyhaline, Mesohaline, and Tidal Freshwater Infaunal similarities and plots calculated from data reported by Weisberg et al. (2008) and Thompson et al. (2012). Plots display clustering overlays at 60% (red dotted lines) and 80% (blue dotted lines) similarity.

Table 11. Species and Taxa Frequently Found by Most Rocky Intertidal Experts to be Useful for Evaluating Community State. High Abundances of Species and Taxa Believed to Reflect Either High or Low Disturbance States.

High Disturbance Indicators	Low Disturbance Indicators
Ulva spp.	Rockweeds (e.g., *Silvetia*, *Hesperophycus*, *Fucus*, *Pelvetiopsis* spp.)
Benthic Diatoms	Owl Limpets (*Lottia gigantea*)
Blue-Green Algae	Black abalone (*Haliotis cracherodii*)
Crustose Red and Brown Algae (Calcified and Non-calcified)	Surfgrass (e.g., *Phyllospadix* spp.)
Small Red Turf-Forming Algae (e.g., *Polysiphonia*, *Ceramium* spp.)	Kelps and Large Brown Seaweeds (e.g., *Egregia menziesii*, *Laminaria* spp., *Eisenia* spp., *Alaria* spp., *Stephanocystis* spp.
Non-native Seaweeds (e.g., *Sargassum muticum, Caulacanthus okamurae*)	
Sand-Influenced Taxa (e.g., *Anthopleura elegantissima, Scytosiphon* spp., *Petalonia* spp.)	
Filamentous Green Algae (e.g., *Chaetomorpha* spp., *Cladophora* spp.)	

Table 12. Importance of Physical Attributes Used to Categorize Rocky Intertidal Sites. Lower Values Indicate Greater Importance. Reported are Means (± 1 SD) of Expert Responses Based on the Following Scale: 1) Provides Critical Information of Primary Importance; 2) Provides Valuable Information of Importance; 3) Provides Information of Value; 4) Provides Information But Used as a Secondary Factor; and 5) Provides Little if any Information; of Limited or No Use. Data Collected Separately for Central and Southern California Rocky Intertidal Assessments but Pooled Because of Consistency in Results With the Exception of the Degree of Wave Exposure, Which Was Considered to be More Important in Categorizing Central California (Mean Response = 1.29) Versus Southern California (2.00) Sites. Qualitative Metrics (e.g., very low, low, moderate, high, very high) Were Used for Parameters Indicated with an Asterisk (*).Degree of Freshwater Influence Was Assessed by Providing Source and Proximity.

Physical Attribute	Mean Response	± 1 SD
*Degree of Sand Influence	1.18	0.37
*Degree of Wave Exposure	1.64	0.72
Primary Substratum Type (e.g., Bedrock, Boulders)	1.93	1.00
*Susceptibility to Substratum Breakout	2.36	0.82
*Substratum Relief	2.54	0.93
Degree of Freshwater Influence	2.64	0.72
*Substratum Character (Degree of Consolidation)	2.71	1.38
Substratum Slope (Degrees)	2.86	1.03
Substratum: Geologic Formation	2.93	1.33
Wave Exposure (Primary Direction)	3.32	0.99
Site Biogeographic Affinity	3.82	1.38
Site Latitude	3.96	1.31

3.6. Physical Environmental Attributes Used to Categorize Communities. Experts focused mostly on three physical environmental variables to characterize community types: the degree of sand influence, the degree of wave exposure, and the nature of the primary substratum (Table 12). Experts almost uniformly agreed that the degree of sand influence was of critical importance in characterizing natural disturbance levels on central and southern California shores. The degree of wave exposure was similarly identified as an important parameter for characterizing central California sites but was thought to be slightly less important for southern California, where wave exposure is generally significantly lower for mainland sites. Primary bench type was the most important of the six substratum parameters for which information was provided. Of the physical environmental parameters, biogeographic affinity and latitude were thought to be less useful in categorizing site communities.

4.0. Discussion

Experts achieved less agreement in evaluating the states of rocky intertidal compared with polyhaline soft-bottom benthic communities (Weisberg et al., 2008; Teixeira et al., 2010); however, agreement was similar, and for central California slightly greater, than in BPJ exercises for tidal freshwater and mesohaline infaunal assemblages (Thompson et al., 2012). Communities inhabiting rocky shores and tidal freshwater and mesohaline soft bottom environments are characterized by greater spatial and temporal variation and subjected to more complex natural disturbance regimes compared with most polyhaline soft bottom habitats, making their states more difficult to interpret using species abundance data (Elliott and Quintino, 2007; Moyle et al., 2010).

State evaluations of rocky intertidal communities present challenges that go beyond those for soft-bottom, infaunal communities. These include: 1) greater spatial heterogeneity in rocky intertidal habitats; 2) less ability of rocky shore experts to assign deviations from an expected natural state to anthropogenic stressors; 3) differences in the characteristics of infauna compared with rocky intertidal macro-organisms; and, 4) the sampling protocols employed to generate the data for the habitat types. In addition, procedural differences, including the number of participants and, in particular, our added evaluation opportunities and discussions, affected comparisons with the BPJ exercises performed by Weisberg et al. (2008), Teixeira et al. (2010), and Thompson et al. (2012).

It has long been recognized that spatial variation in rocky intertidal communities is high due to environmental features such as substratum characteristics, tidal position, and exposure to wave energy, sand influence, and freshwater input. Natural disturbances, for example from storm waves or extreme low tide desiccation events, also generate temporal variation in community structure over a range of spatial scales. This variability makes it difficult to detect anthropogenic disturbance signatures, even when known, and requires experts to integrate small and large scale variation in habitat features together with spatial and temporal effects of multiple, frequent natural disturbances in making state evaluations. These difficulties are enhanced when sources and magnitudes of anthropogenic stress are unknown or not provided.

Spatial and temporal environmental variability also create challenges in evaluating the states of soft bottom communities even though the biological signatures of common anthropogenic perturbations (e.g., organic and other forms of sediment contamination) are well known and generally differ from major natural disturbances (Pearson and Rosenberg, 1978; Bilyard, 1987; Díaz and Rosenberg, 1995; Borja et al., 2003; Díaz et al., 2004; Marques et al., 2009; Pinto et al., 2009). This is particularly true in tidal freshwater and estuarine habitats where small-scale environmental gradients and the high frequency of natural disturbance make it difficult to distinguish anthropogenic from natural impacts (Dauvin, 2007; Elliott and Quintino, 2007). Development of biologically-based indices for estuaries has lagged because of the confounding effects of multiple natural environmental stressors (Elliott and Quintino, 2007). By contrast, environmental conditions in polyhaline habitats are generally more stable and vary mostly over large spatial scales (Ranasinghe et al., 2012); thus, common anthropogenic disturbances are easier to distinguish and detect. For polyhaline communities, where experts achieved strong agreement, the observed variation in state evaluations was attributed mostly to differences in views on the relative importance of accepted biological indicators of anthropogenic perturbation not the signatures themselves (Weisberg et al., 2008; Teixeira et al., 2010). In tidal freshwater and mesohaline communities, where there was less expert agreement, variation in state evaluations was attributed to the confounding effects of high spatial and temporal environmental variability and difficulties in distinguishing anthropogenic from natural biological signatures under these conditions (Thompson et al., 2012).

Species types and abundances provide the biological information for BPJ determinations of community state. The types of species used in our rocky shore exercise consisted of macro-invertebrates and macrophytes instead of smaller, infaunal organisms. Besides being larger in size, these macro-organisms generally have longer life spans, slower turnover rates, and their populations often respond more slowly to disturbances. Infauna are also more directly exposed to sediment contamination and other anthropogenic impacts and species distributions and abundances are known to be strong indicators of responses to such environmental stressors (Bilyard, 1987; Díaz et al., 2004, Bay et al., 2007). However, it is difficult to identify consistent biological indicators of most forms of anthropogenic perturbation on rocky shores because there is high natural variation among sites in the abundances and even the presence of many species (Foster, 1990, Zabin et al., 2012) and anthropogenic impacts are added to the multiple effects of highly fluctuating and often stressful natural disturbances. In addition, in rocky intertidal habitats the biological responses to many of these perturbations are similar to naturally-occurring stressors. Thus, it is difficult to distinguish anthropogenic from natural changes in community composition, particularly where the sources or types of anthropogenic perturbations are unknown and their effects are not extreme.

BPJ comparisons of rocky intertidal macro-organism communities with infaunal assemblages also are complicated by differences in sampling procedures. In BPJ infaunal exercises, the biological data representing each analyzed site were obtained from a single core sample whose contents were sorted and identified in the laboratory. By contrast, the rocky intertidal community data were derived from multiple, field-identified samples distributed over a similar range of tidal elevations, an approach that likely generates greater average richness and more consistency in species content among site assemblages. For example, infaunal richness averaged 30, 14, and 9 taxa for west coast polyhaline, mesohaline, and tidal freshwater communities and ranged from 1 to 73, 3 to 25, and 3 to 24, respectively for these three habitat types (Weisberg et al., 2008; Thompson et al., 2012). These values were very different from those for the examined rocky intertidal communities, where richness averaged 53 (central California) and 45 (southern California) and ranged from 30 to 78 and 34 to 60 taxa. This is important because infaunal experts consistently used low community richness as an indicator of strong anthropogenic impacts in their evaluations (Weisberg et al., 2008; Texeira et al. 2010; Thompson et al., 2012).

BPJ state evaluations of rocky shore communities are affected by greater difficulty in associating biological with environmental variables. Structuring environmental parameters (e.g., grain size, depth, sediment contaminants) are strongly linked to infaunal community structure in each core sample. However, on rocky shores, key physical environmental variables (e.g., substratum characteristics, wave action, sand influence) are often not easily associated with site-level species distributions and abundances; these variables typically represent conditions averaged over coarse spatial scales.

Unlike the BPJ infaunal exercises, we did not complete rocky intertidal state evaluations in a single trial. Although the data sets employed were similar, rocky shore experts were uncomfortable in making their final state evaluations after the first two BPJ trials because of differing interpretations of the scoring system and difficulties in relating environmental to biological data. After learning the identities of the sites, experts also expressed concern that sites subjected to high levels of anthropogenic impact were poorly represented in the first central California trial. They also had trouble determining community states using data from only a single point in time without knowledge of whether an extreme natural disturbance event (e.g., storm) had occurred prior to sampling. As expected, small increases in the level of expert agreement were reached from the first to the last trial, likely due to changes in the evaluation scale and because more information became available. However, the increase in the level of expert agreement was less as was the average agreement achieved for southern California sites, most likely because sites were exposed to greater geographic variation in ocean conditions and many experts were less familiar with this region.

Difficulties in achieving uniform understanding of the evaluation scale are common in BPJ exercises (Bay et al., 2007; Weisberg, personal communication). Because of the high spatial and temporal variability of rocky shores, experts wanted more information on site features before making final determinations and sought changes in application of the disturbance scale to incorporate site-to-site uniqueness in natural disturbance regimes. As expected, the addition of urban southern California sites added more urban and potentially anthropogenically-disturbed communities to the BPJ exercise. The range of anthropogenic impacts represented by the rocky intertidal data sets, however, were still limited compared with infaunal BPJ data sets (Weisberg et al., 2008; Thompson et al., 2012) where a wide range of contamination conditions were selected *a priori* using Long and MacDonald's (1998) mean Effects Range-Median quotient (mERMq). An analogous single measure of perturbation to the mERMq is not available for rocky shore communities to ensure the representation of a wide range of anthropogenically disturbed states. Less expert agreement is likely when assemblages are located nearer the center of the disturbance scale and highly impacted communities are poorly represented (Borja et al., 2009; Texeira et al., 2010) as was the case in our rocky intertidal BPJ exercises.

Experts agreed that evaluations of the ecological states of rocky intertidal and other spatially and temporally variable communities are strongly challenged when judgments are based only on a single snapshot in time because a single sampling point does not allow evaluations to take into account recent, major natural disturbance or recruitment events or to be made in the context of the range of expected states for that site. A range of community states is the norm for any site, and in our exercise deviation from an expected state was the basis for distinguishing anthropogenic or major natural disturbances. However, although available, we did not include data sets obtained over multiple years in this exercise because in practice few rocky intertidal site sampling programs have been carried out over sufficient time to enable an expected range of states to be determined.

Although there is high variation among rocky intertidal sites in species abundances and even species presence (Foster, 1990, Zabin et al., 2012), most experts relied on certain species types in making their final evaluations, particularly macrophytes with functional characteristics related to morphology (*sensu* Littler and Littler, 1980) and sessile macro-invertebrates; most experts placed less emphasis on mobile macro-invertebrates such as littorine (*Littorina* spp.) snails, whose distributions and abundances are often highly variable and difficult to interpret. Smaller, short-lived, morphologically simple macroalgae with fast growth rates and high reproductive outputs (e.g., *Ulva* spp. and small, red and green turf-forming and filamentous algae) were often used as possible indicators of recent or continuous disturbance. By contrast, slower growing, longer-lived, and morphologically complex rockweeds and lower-shore, large brown seaweeds were thought to be possible indicators of more stable, less impacted communities. Macrophytes have been used in the past to characterize the ecological status of coastal communities (Littler and Littler, 1981, 1984; Orfanidis et al., 2001; Ballesteros et al., 2007; Juanes et al., 2008; Schramm, 1999), and in attempts to develop an ecological evaluation index within the European Water Directive (Orfanidis et al., 2001, 2003; Panayotidis et al., 2004). However, abundances of macrophyte types alone cannot differentiate natural from anthropogenic disturbance. For example, high abundances of *Ulva* spp. and other smaller, opportunistic algae not only characterize sewage-impacted rocky shores (e.g., Littler and Murray, 1975) but also shores disturbed by naturally occurring sand (Littler et al., 1991; Murray and Bray, 1993; Airoldi, 2003) and boulder movements (Sousa, 1979, 1980). Large, conspicuous macro-invertebrates, such as owl limpets and black abalone (Keough et al., 1993; Miller and Lawrenz-Miller, 1993; Addessi, 1994; Sagarin et al., 2006), and intertidal rockweeds (Bokn and Lein, 1978; Bokn, 1979; Vogt and Schramm, 1991; Rodriguez-Prieto and Polo, 1996; Oliveira and Qi, 2003) are affected by human activities and reduced abundances of these taxa were also considered by experts to be indicators of anthropogenic disturbance as was the presence of non-indigenous species based on the premise that disturbed habitats are more susceptible to invasion (Dukes and Mooney, 1999; Byers, 2002). Abundances of environmentally sensitive and tolerant taxa were identified as important criteria for detecting anthropogenic disturbance in BPJ infaunal exercises (Weisberg et al., 2008; Teixeira et al., 2010; Thompson et al., 2012) and are regarded as strong indicators of community state in soft bottom habitats (Borja et al., 2000; Muxika et al., 2005; Dauvin, 2007), particularly where focus is on a single type of stressor (e.g., organic sediment contamination). Unfortunately, knowledge of tolerant and sensitive indicator taxa is less developed for rocky intertidal communities, where multiple and often unknown stressors are the norm, and there is no consensus on either a universal disturbance paradigm or consistent and reliable biological indicators of most types of anthropogenic stress (Murray et al., 2006).

Managers are often asked to make ecological evaluations of coastal communities that take into account the effects of multiple natural and anthropogenic stressors. BPJ exercises can inform such efforts by summarizing expert opinion on the states of ecological communities over a wide range of available data, determining the degree of expert consensus, helping to identify key biological indicators, and by calibrating and evaluating index performance where an index exists (Ranasinghe, et al., 2008, Weisberg et al., 2008; Teixeira et al., 2010; Thompson et al., 2012). In this BPJ exercise, rocky intertidal experts failed to achieve the level of agreement reported for experts working with polyhaline infaunal communities; however, they did as well as infaunal experts evaluating mesohaline assemblages and better than experts evaluating tidal freshwater assemblages. Interestingly, despite the relatively low agreement achieved among experts for mesohaline soft-bottom communities, indices have been developed and evaluated by

BPJ exercises that are considered good enough to be part of the regulatory assessment process in California and elsewhere (Beegan and Bay, 2012). This study underscores the difficulties in distinguishing deviations from an expected natural state on rocky shores working from “one-off” species abundance data and physical site descriptions without information on the sources and magnitudes of anthropogenic perturbation. Such difficulties are likely in other habitats subjected to multiple disturbances, and high spatial and temporal variation, particularly where biological responses to natural and anthropogenic stressors are often similar and difficult to distinguish without more information. Hence, if the goal is to use BPJ exercises to identify anthropogenically-impacted or strongly disturbed sites for rocky shores and similar habitats, expert opinions must be subjected to rigorous testing to firmly establish links between community states and known environmental stressors.

5.0 Acknowledgments

This exercise grew out of discussions on data needed to monitor and evaluate the ecological states of west coast rocky intertidal communities initiated by participants in the Multi-Agency Rocky Intertidal Network (MARINe), a partnership of agencies, universities and private groups committed to determining the health of rocky intertidal habitats. The U.S. Bureau of Ocean Energy Management (BOEM), a primary sponsor of MARINe, the University of Southern California (USC) Sea Grant program (National Oceanic and Atmospheric Administration (NOAA), the Santa Monica Bay Restoration Commission (SMBRC), and the Southern California Coastal Water Research Project Authority (SCCWRP) provided fiscal and logistic support for this project. Pete Raimondi’s group at the University of California, Santa Cruz (UCSC) provided the data sets, which were obtained with funding from MARINe, the

Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO), and the Division of Water Quality (Ocean Unit) of California's State Water Control Board (SWCB). We are particularly grateful to Mary Elaine Helix (BOEM) and Guangyu Wang (SMBRC) for providing travel funds for expert participants for post-exercise workshops and Linda Duguay, Phyllis Grifman, and Susan Zaleski (USC Sea Grant) for arranging for logistics and underwriting costs of an initial workshop at the Catalina Island Wrigley Institute of Environmental Studies (WIES). We also thank Liz Whiteman (California Ocean Science Trust (OST) and the OST's Marine Protected Area Monitoring Enterprise) and Dominic Gregorio (Ocean Unit of the Division of Water Quality, SWCB) for their interest and inputs. SCCWRP sponsored meetings and webinars essential to the completion of this project and Rani Gaddam (UCSC), Christy Bell (UCSC), Melissa Miner (UCSC), and Carol Blanchette (Marine Science Institute; University of California Santa Barbara) provided assistance in collecting and compiling data sets and in developing physical descriptions of rocky intertidal sites. Professor K.R. (Bob) Clarke (PRIMER-E Ltd) provided advice in data analysis. The views expressed herein are those of the authors and do not necessarily reflect the views of BOEM, NOAA, OST, SMBRC, or SWCB, or any of their subagencies.

6.0. Literature Cited

Addessi, L., 1994. Human disturbance and long-term changes on a rocky intertidal community. Ecological Applications 4, 786-797.

Airoldi, L., 2003. The effects of sedimentation on rocky coast assemblages. Oceanography and Maine Biology, An Annual Review 41, 161-236.

Bard, S.M., 1998. A biological index to predict pulp mill pollution levels. Water Environment Research. 70, 108-122.

Ballesteros, E., Torras, X., Pinedo, S., García, M., Mangialajo, L., Torres, M., 2007. A new methodology based on littoral community cartography dominated by macroalgae for the implementation of the European water framework directive. Marine Pollution Bulletin 55, 172-180.

Bay, S., Berry, W., Chapman, P.M., Fairey, R., Gries, T., Long, E., MacDonald, D., Weisberg, S.B., 2007. Evaluating consistency of best professional judgment in the application of a multiple lines of evidence sediment quality triad. Integrated Environmental Assessment and Management 3, 491-497.

Beegan, C., Bay, S.M., 2012. Transitioning sediment quality assessment into regulations: challenges and solutions in implementing California's sediment quality objectives. Integrated Environmental Assessment and Management 8, 586-588.

Bilyard, G.R., 1987. The value of benthic infauna in marine pollution monitoring studies. Marine Pollution Bulletin 11, 581-585.

Bokn, T., 1979. Use of benthic algae classes as indicators for eutrophication in estuarine and marine waters. In: Hytteborn, H. (Ed.), The Use of Ecological Variables in Environmental Monitoring. Naturvardsverket, report 1151.

Bokn, T., Lein, T.E., 1978. Long-term changes in fucoid association of the inner Oslofjord, Norway. Norwegian Journal of Botany 25, 9-14.

Borja, A., Miles, A., Occhipinti-Ambrogi, A., Berg, T., 2009. Current status of macroinvertebrate methods used for assessing the quality of European marine waters: implementing the Water Framework Directive. Hydrobiologia 63, 181-196.

Borja, A., Muxika, I., Franco, J., 2003. The application of a Marine Biotic Index to different impact sources affecting soft-bottom benthic communities along European coasts. Marine Pollution Bulletin 46, 835-845.

Borja, A., Franco, J., Perez, V., 2000. A marine biotic index to establish the ecological quality of soft bottom benthos within European estuarine and coastal environments. Marine Pollution Bulletin 40, 1100-1114.

Byers, J.E., 2002. Impact of non-indigenous species on natives enhanced by anthropogenic alteration of selection regimes. Oikos 97, 449-458.

Clarke, K.R., 1993. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology 18, 117-143.

Clarke, K.R., Gorley, R.N., 2006. PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth, U.K.

Dauvin, J.C., 2007. Paradox of estuarine quality: benthic indicators and indices, consensus or debate for the future. Marine Pollution Bulletin 55, 271-281.

Díaz, R.J., Rosenberg, R., 1995. Marine benthic hypoxia: a review of its ecological effects and the behavioural responses of benthic macrofauna. Oceanography and Marine Biology: An Annual Review 33, 245-303.

Díaz, R.J., Solan, M., Valente, R.M., 2004. A review of approaches for classifying benthic habitats and evaluating habitat quality. Journal of Environmental Management 73, 165-181.

Dukes, J.S., Mooney, H.A., 1999. Does global change increase the success of biological invaders. Trends in Ecology and Evolution 14, 135-139.

EC, 2000. Directive 2000/60/EC of the European Parliament and of the Council: establishing a framework for community action in the field of water policy. Official Journal of the European Communities (OJ L327) 43, 1–73.

Elliott, M., Quintino, V.M., 2007. The estuarine quality paradox, environmental homeostasis and the difficulty of detecting anthropogenic stress in naturally stressed areas. Marine Pollution Bulletin 54, 640-645.

Foster, M.S., 1990. Organization of macroalgal assemblages in the Northeast Pacific: the assumption of homogeneity and the illusion of generality. Hydrobiologia 192, 21-33.

Juanes, J.A., Guinda, X., Puente, A., Revilla, J.A., 2008. Macroalgae, a suitable indicator of the ecological status of coastal rocky communities in the NE Atlantic. Ecological Indicators 8, 351-359.

Keough, M.J., Quinn, G.P., King, A., 1993. Correlations between human collecting and intertidal mollusc populations on rocky shores. Conservation Biology 7, 378-390.

Littler, M.M., Littler, D.S., 1980. The evolution of thallus form and survival strategies in benthic marine macroalgae: field and laboratory tests of a functional form model. American Naturalist 116, 25-44.

Littler, M.M., Littler, D.S., 1981. Intertidal macrophyte communities from Pacific Baja California and the upper Gulf of California: relatively constant vs. environmentally fluctuating systems. Marine Ecology Progress Series 4, 145-158.

Littler, M.M., Littler, D.S., 1984. Relationships between macroalgal functional form groups and substrata stability in a subtropical rocky-intertidal ecosystem. Journal of Experimental Marine Biology and Ecology 74, 13-34.

Littler, M.M., Murray, S.N., l975. Impact of sewage on the distribution, abundance and community structure of rocky intertidal macro-organisms. Marine Biology 30, 277-29l.

Littler, M.M., Littler, D.S., Murray, S.N., Seapy, R.R., 1991. Southern California rocky intertidal ecosystems. In: Mathieson, A.C., Nienhuis, P.H. [Eds.], Ecosystems of the World, Volume 24. Intertidal and Littoral Ecosystems, Elsevier, Amsterdam, 273-96.

Long, E.R., MacDonald, D.D., 1998. Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems. Human Ecological Risk Assessment 4, 1019-1039.

Marques, J.C., Salas F., Patrício, J., Teixeira, H., Neto, J.M., 2009. Ecological Indicators for Coastal and Estuarine Environmental Assessment - a user guide. WIT Press, Southampton.

Meyer, M.A., Booker. J.M., 2001. Eliciting and analyzing expert judgment. A practical guide. Society for Industrial and Applied Mathematics. Philadelphia Pennsylvania.

Miller, A.C., Lawrenz-Miller, S.E., 1993. Long-term trends in black abalone, *Haliotis cracherodii* Leach, 1814, populations along the Palos Verdes Peninsula, California. Journal of Shellfish Research 12, 195-200.

Moyle, P.B., Bennett, W.A., Fleenor, W.E., Lund, J.R., 2010. Habitat variability and complexity in the upper San Francisco Estuary. San Francisco Estuary and Watershed Science 8, 1-24.

Murray, S.N., Ambrose, R.F., Dethier, M.N., 2006. Monitoring Rocky Shores. University of California Press, Berkeley. 220 pp.

Murray, S.N., Bray, R.N., 1993. Benthic Macrophytes In: Dailey, M.D., Reish, D.J., Anderson, J.W., (Eds.), Ecology of the Southern California Bight: A Synthesis and Interpretation. University of California Press, Berkeley, 304-368.

Muxika, I., Borja, A., Bonne, W., 2005. The suitability of the marine biotic index (AMBI) to new impact sources along the European coasts. Ecological Indicators 5, 19-31.

Oliveira, E.C., Qi, Y., 2003. Decadal changes in a polluted bay as seen from its seaweed flora: the case of Santos Bay in Brazil. Ambio 32, 403-405

Orfanidis, S., Panayotidis, P., Stamatis, N., 2001. Ecological evaluation of transitional and coastal waters: a marine benthic macrophytes based model. Mediterranean Marine Science 2, 45-65.

Orfanidis, S., Panayotidis, P., Stamatis, N., 2003. An insight to the ecological evaluation index (EEI). Ecological Indicators 3, 27-33.

Panayotidis, P., Montesanto, B., Orfanidis, S., 2004. Use of low budget monitoring of macroalgae to implemement the European Water Framework Directive. Journal of Applied Phycology 16, 49-59.

Pearson, T.H., Rosenberg, R., 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Oceanography and Marine Biology: An Annual Review 16, 229-311.

Pinedo, S., García, M., Satta, M.P., de Torres, M., Ballesteros, E., 2007. Rocky-shore communities as indicators of water quality: A case study in the Northwestern Mediterranean. Marine Pollution Bulletin 55,126–135.

Pinto, R., Patricio, J., Baeta, A., Fath, B.D., Aneto, J.M., and Marques, J.C., 2009. Review and evaluation of estuarine biotic indices to assess benthic condition. Ecological Indicators 9, 1-25.

Ranasinghe, J.A., Weisberg, S.B., Smith, R.W., Montagne, D.E., Thompson, B., Oakden, J.M., Huff, D.D., Cadien, D.B., Velarde, R.G., Ritter, K.J., 2008. Calibration and evaluation of five indicators of benthic community condition in two California bay and estuary habitats. Marine Pollution Bulletin 59, 5-13.

Ranasinghe, J.A., Welch, K.I., Slattery, P.N., Montagne, D.E., Huff, D.D., Lee II, H., Hyland, J.L., Thompson, B., Weisberg, S.B., Oakden, J.M., Cadien, D.B., Velarde, R.G., 2012. Habitat-related benthic macrofaunal assemblages of bays and estuaries of the western United States. Integrated Environmental Assessment and Management 8, 638-648.

Revilla, M., Franco, J., Bald, J., Laza, A., Seoane, S., Valencia, V., 2009. Assessment of the phytoplankton ecological status in the Basque coast (northern Spain) according to the European Water Framework Directive. Journal of Sea Research 61, 60-67.

Rodriguez-Prieto, C., Polo, L., 1996. Effects of sewage pollution in the structure and dynamics of the community of *Cystoseira mediterranea* (Fucales, Phaeophyceae). Scientia Marina 60, 253-263.

Sagarin, R.D., Ambrose, R.F., Becker, B.J., Engle, J.M., Kido, J., Lee, S.F., Miner, C.M., Murray, S.N., Raimondi, P.T., Richards, D.V., Roe, C., 2006. Ecological Impacts on the limpet *Lottia gigantea* populations: human pressure over a broad scale on islands and mainland intertidal zones. Marine Biology 150, 399-413.

Schoch, G.C., Dethier, M.N., 1996. Scaling up: the statistical linkage between organismal abundance and geomorphology on rocky intertidal shorelines. Journal of Experimental Marine Biology and Ecology 201, 37–72.

Schramm, W., 1999. Factors influencing seaweed responses to eutrophication: some results from EU-project EUMAC. Journal of Applied Phycology 11, 69-78.

Selig, U., Eggert, A., Schories, D., Schubert, M., Blumel, C., Schubert, H., 2007. Ecological classification of macroalgae and angiosperm communities of inner coastal waters in the southern Baltic Sea. Ecological Indicators 7, 665-678.

Sfriso, A., Facca, C., Ghetti, P., 2009. Validation of the Macrophyte Quality Index (MaQI) set up to assess the ecological status of Italian marine transitional environments. Hydrobiologia 617, 117-141.

Sousa, W.P., 1979. Experimental investigations of disturbance and ecological succession in a rocky intertidal algal community. Ecological Monographs 49, 227-254.

Sousa, W.P., 1980. The responses of a community to disturbance: the importance of successional age and species life histories. Oecologia 45, 72-81.

Teixeira, H., Borja, A., Weisberg, S.B., Ranasinghe, J.A., Cadien, D.B., Dauer, D.M., Dauvin, J.C., Degraer, S., Diaz, R.J., Grémare, A., Karakassis, I., Llansó, R.J., Lovell, L.L., Marques, J.C., Montagne, D.E., Occhipinti-Ambrogi, A., Rosenberg, R., Sardá, R., Shaffner, L.C., Velarde, R.G., 2010. Assessing coastal benthic macrofauna community condition using best professional judgment - Developing consensus across North America and Europe. Marine Pollution Bulletin 60, 589-600.

Thompson, B., Weisberg, S.B., Melwani, A., Lowe, S., Ranasinghe, J.A., Cadien, D.B., Dauer, D.M., Díaz, R.J., Fields, W., Kellogg, M., Montagne, D.E., Ode, P.R., Reish, D.J., Slattery, P.N., 2012. Low levels of agreement among experts using best professional judgment to assess benthic condition in the San Francisco Estuary and Delta. Ecological Indicators 12, 167-173.

Vogt, H., Schramm, W., 1991. Conspicuous decline of *Fucus* in Kiel Bay (Western Baltic): what are the causes? Marine Ecology Progress Series 69, 189-194.

Weisberg, S.B., Ranasinghe, J.A., Dauer, D.M., Schaffner, L.C., Diaz, R.J., Frithsen, J.B., 1997. An estuarine benthic index of biotic integrity (B-IBI) for Chesapeake Bay. Estuaries 20:149-158.

Weisberg, S.B.,Thompson, B., Ranasinghe, J.A., Montagne, D.E., Cadien, D.B., Dauer, D.M., Diener, D., Oliver, J., Reish, D.J., Velarde, R.G., Word,J.Q., 2008. The level of agreement among experts applying best professional judgment to assess the condition of benthic infaunal communities. Ecological Indicators 8, 389-394.

Wilkinson, M., Wood, P., Wells, E., Scanlan, C., 2007. Using attached macroalgae to assess ecological status of British estuaries for the European Water Framework Directive. Marine Pollution Bulletin 55, 136-150.

Zabin, C.J., Danner, E.M., Baumgartner, E.P., Spafford, D., Miller, K.A., Pearse, J.S., 2012. A comparison of intertidal species richness and composition between Central California and Oahu, Hawaii. Marine Ecology, 34, 131-156.