Journal of Environmental Management 268 (2020) 110676

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: http://www.elsevier.com/locate/jenvman

Research article

Eliciting expert judgment to inform management of diverse oyster resources for multiple ecosystem services

Daniel R. Petrolia [a,*], Frederick Nyanzu [b], Just Cebrian [c], Ardian Harri [a], Jamie Amato [d,e], William C. Walton [f]

[a] Department of Agricultural Economics, Mississippi State University, Box 5187, Mississippi State, MS, 39762, USA
[b] Department of Agricultural & Consumer Economics, University of Illinois, 326 Mumford Hall, Urbana, IL, 61801, USA
[c] Northern Gulf Institute, Mississippi State University, 1021 Balch Blvd, Stennis Space Center, MS, 39529, USA
[d] Department of Marine Sciences, University of South Alabama, 25 Life Sciences Bldg, Mobile, AL, 36688, USA
[e] Dauphin Island Sea Lab, 101 Bienville Blvd, Dauphin Island, AL, 36528, USA
[f] Auburn University Shellfish Laboratory, Auburn University, 150 Agassiz St., Dauphin Island, AL, 36528, USA

ARTICLE INFO

Keywords:
Coastal
Erosion
Expert elicitation
Marine
Off-bottom aquaculture
Restored reef

ABSTRACT

This study presents the most comprehensive set of ecosystem service provision estimates for diverse oyster-based resources to date. We use expert elicitation methods to derive estimates of five ecosystem services provided by oysters: oyster harvest (as indicated by oyster density), improved water quality (net nitrogen assimilation), shoreline protection (net erosion), and other fish habitat (blue crab and red drum density). Distributions are estimated for three distinct resources: on-bottom production, off-bottom farms, and non-harvested restoration/conservation efforts, under twelve distinct scenarios according to varying environmental conditions (eutrophication, sedimentation, and salinity regimes). Our expert-derived estimates of ecosystem services provide useful comparisons across oyster resources of both expected ecosystem service delivery levels and the amount of variation in those levels. These estimates bridge an information gap regarding relative performance of diverse oyster resources along multiple dimensions and should serve as a useful guide for resource managers facing competing interests.

1. Introduction

It is well-documented that oysters provide a variety of ecosystem services (Alleway et al., 2019; Fodrie et al., 2017; Grabowski et al., 2012; Higgins et al., 2011; Humphries and La Peyre, 2015; Interis and Petrolia, 2016; Kellogg et al., 2014; Meyer et al., 1997; Piazza et al., 2005; Piehler and Smyth, 2011; Scyphers et al., 2011; Smyth et al., 2013). Nevertheless, there remains a gap in the literature regarding the distribution of the levels of services delivered by oysters and oyster resources, particularly the differences that exist among harvested and non-harvested oyster resources. Here we use the term *oyster resources* to distinguish oyster production methods, grouped broadly into three categories: on-bottom production (traditional bottom leases and commercially harvested oyster beds), off-bottom farming (containers where oysters are kept off the bottom, including cases where the container itself sits on the bottom), and restoration/conservation efforts (living shorelines and restored reefs with the intention of no-harvest). Hereafter, we refer to these as *traditional*, *off-bottom*, and *restored* for convenience (see Table 1 for summary of key terms and descriptions). Although it is well documented that *restored* yields a suite of ecosystem services, those derived from *traditional* and *off-bottom* are less known. This hinders our understanding of how diverse oyster resources differ in their ecosystem services and, thus, limits our capacity to manage these valuable resources. Indeed, there may exist substantial differences in the type and extent of ecosystem services provided by oyster resources due, in part, to their specific responses to changing environmental conditions. For example, under high-salinity scenarios along the coast of the Gulf of Mexico, *restored* and *traditional* struggle to deliver ecosystem services due to mortalities stemming from predation and disease. On the other hand, ecosystem service delivery by *off-bottom* does not appear to diminish as they avoid the losses suffered by oysters on the bottom (e.g., Mann and Powell, 2007).

The gap in information regarding ecosystem services provided by harvested oyster resources, and how it compares with those provided by

* Corresponding author.
E-mail address: d.petrolia@msstate.edu (D.R. Petrolia).

https://doi.org/10.1016/j.jenvman.2020.110676
Received 18 October 2019; Received in revised form 28 April 2020; Accepted 29 April 2020
Available online 14 May 2020
0301-4797/

non-harvested resources, is particularly acute given the demand for this kind of information by scientists, managers, and regulators, especially when faced with decisions about multiple oyster resources and potential user conflicts (Baggett et al., 2015; Sutton-Grier and Bamford, 2015). Grabowski et al. (2012) note that "decision-support tools that allow location-specific value estimates would be of great value in prioritizing restoration projects to enhance the value of the services that they provide" (p. 907). Similarly, Alleway et al. (2019) note the importance of "increasing recognition, understanding, and accounting of ecosystem service provision by mariculture" (p. 59). Indeed, one of the main messages of Grabowski et al. (2012) is that different oyster resources should be considered when evaluating oyster-derived ecosystem services, a topic that we directly address in our study.

The purpose of this research is to generate reliable estimates of ecosystem service levels and distributions of direct relevance to the management of diverse oyster resources using expert elicitation methods. Toward this purpose, the work answers two specific questions: 1) *Do ecosystem services differ across oyster resources?* and 2) *Do these estimates vary based on region and elicitation method?* We provide an answer to the first question by generating novel estimates of five intermediate ecosystem services – oyster density, net nitrogen assimilation, reduced erosion, blue crab density, and red drum density – that can be mapped, under reasonable assumptions, to final ecosystem services, respectively –oyster harvest (for harvested resources only), improved water quality, reduced erosion, blue crab habitat, and red drum habitat. We provide these estimates for all three of the aforementioned oyster resources under 12 different scenarios of environmental conditions according to eutrophication, sedimentation, and salinity levels (see Table 1).

We provide an answer to the second question using a sample comprised of experts from two U.S. oyster-producing regions: the Atlantic and Gulf Coasts, and an experimental design that includes two different approaches for eliciting expert knowledge. Recent research, discussed below, has found that the elicitation method can affect responses. Our estimates are based on data collected from an extensive questionnaire designed to elicit oyster expert knowledge, an approach "used widely in the science and practice of conservation because of the complexity of problems, relative lack of data, and the imminent nature of many conservation decisions" (p. 29, Martin et al., 2012). A recent example is the case of ranking threats to endangered species (Donlan et al., 2010), and other examples in the literature include impacts of plastic pollution on marine wildlife (Wilcox et al., 2016), threatened species (McBride et al., 2012), and climate change impacts (Fuentes and Cinner, 2010) and coastal modification (Nelson Sella et al., 2019) on sea turtle nesting grounds.

We present our expert-derived estimates as an alternative to, but certainly not as a complete or comprehensive substitute for, yet-unavailable new studies providing original empirical data. Despite the

Table 1
Key terms and descriptions.

Oyster Resources

Resource	Description
Traditional	natural oyster beds/reefs and/or traditional on-bottom plantings
Off-bottom	containers where oysters are kept off the bottom, including cases where the container itself sits on the bottom
Restored	living shorelines, restored reefs with the intention of no-harvest

Ecosystem Services

Service	Units	As indicator for final service
Oyster density	#/m^2	Oyster harvest
Net N assimilated	g N/m^2	Water quality
Reduced erosion	m shoreline/m^2	Reduced erosion
Blue crab density	#/m^2	Blue crab habitat
Red drum density	#/m^2	Red drum habitat

Environmental Conditions

Category	Levels	Description
Nutrients-Oxygen	Mesotrophic-normoxic	medium nutrient concentrations: < 10 µ mols of dissolved inorganic nitrogen per liter, < 25 µg of chlorophyll-a per liter; well-oxygenated: ≥ 4 mg of oxygen per liter
	Eutrophic-hypoxic	high nutrient concentrations: ≥ 10 µ mols of dissolved inorganic nitrogen per liter, < 25 µg of chlorophyll-a per liter; low oxygen: < 4 mg of oxygen per liter
Sedimentation	Normal	normal
	High	high sedimentation from sources such as river inputs, eroding shorelines, and/or uncontained dredged sediments
Salinity	Low	<10 ppt
	Medium	10–20 ppt
	High	>20 ppt

Environmental Scenarios

Scenario	Nutrients-Oxygen	Sedimentation	Salinity
MNL	mesotrophic-normoxic	normal	low
MNM	mesotrophic-normoxic	normal	medium
MNH	mesotrophic-normoxic	normal	high
MHL	mesotrophic-normoxic	high	low
MHM	mesotrophic-normoxic	high	medium
MHH	mesotrophic-normoxic	high	high
ENL	eutrophic-hypoxic	normal	low
ENM	eutrophic-hypoxic	normal	medium
ENH	eutrophic-hypoxic	normal	high
EHL	eutrophic-hypoxic	high	low
EHM	eutrophic-hypoxic	high	medium
EHH	eutrophic-hypoxic	high	high

recognition that our study cannot replace the rigor of much needed studies providing empirical data, we contend it still presents the most comprehensive set of ecosystem service estimates for diverse (harvested and non-harvested) oyster resources to date.

2. Elicitation of expert knowledge

Martin et al. (2012) define expert knowledge as "substantive information on a particular topic that is not widely known by others; " they define an expert as "someone who holds this knowledge and who is often deferred to in its interpretation; " and they define expert judgments as "predictions by experts of what may happen in a particular context" (p. 29). Elicitation of expert knowledge has been used in a wide variety of settings where the demand for information exceeds its supply. The earliest documented instance of a structural method for eliciting expert knowledge appears to be Dalkey and Helmer (1963), who developed the Delphi method to determine the optimal number of atomic bombs to be manufactured during the Cold War. The method was developed further in Dalkey (1967, 1969). Subsequent uses include Gordon and Helmer (1964), to explore long-term trends in scientific research and forecasted impacts on society, with more recent applications including pharma-economics (Evans, 1997), knowledge management (Holsapple and Joshi, 2002), and e-commerce diffusion (Okoli and Pawlowski, 2004). Applications specific to the environment and conservation efforts include studies on discount rates for future climate-related damages (Weitzman, 2001), ecosystem-based spatial planning (Scolozzi et al., 2012), and ecosystem service valuation (Curtis, 2004), including valuation of Amazon rainforest protection (Strand et al., 2017). There are a variety of other structural methods for eliciting expert knowledge, including Martin et al. (2012); the IDEA protocol of Hemming et al. (2018); O'Hagan et al. (2006), which is specifically focused on eliciting probabilities; and Veen et al. (2017).

The IDEA protocol (the acronym stands for "Investigate, Discuss, Estimate, and Aggregate") is a similar procedure, but with a key difference: the experts of a Delphi exercise work in isolation, whereas the IDEA protocol explicitly includes an open discussion stage between the two elicitation stages. A summary of the basic steps is as follows. A diverse group of experts is recruited to answer questions with probabilistic or quantitative responses. The experts are asked to first *Investigate* the questions and to clarify their meanings, and then to provide their private, individual best guess point estimates and associated credible intervals (Round 1). The experts receive feedback on their estimates in relation to other experts. With assistance of a facilitator, the experts are encouraged to *Discuss* the results, resolve different interpretations of the questions, cross-examine reasoning and evidence, and then provide a second and final private *Estimate* (Round 2). The individual estimates are then combined using mathematical *Aggregation*. Hemming et al. (2018) write that the purpose of discussion in the IDEA protocol is not to reach consensus but to resolve linguistic ambiguity, promote critical thinking, and to share evidence. Hanea et al. (2018) find that it can increase response accuracy. Strand et al. (2017), a recent example of an application of the Delphi method, summarize the key elements of the Delphi method as: (a) anonymous responses by experts to multiple rounds of formal questionnaires; (b) an exercise incorporating iterative, controlled feedback with respect to the information provided at each round; and (c) statistical summary of the group's responses. The present paper follows the Delphi method, and incorporates the open discussion component of the IDEA protocol as part of the experimental design.

Consistent with the above methodologies, our methodology includes an experimental treatment, wherein there are two groups of experts. The first group, which we refer to as the "Isolated Group", completes the entire process online, working in isolation from one another as in Delphi. The second group, which we refer to as the "Discussion Group", completes the initial elicitation round online, exactly as the first group does. However, this group then meets in an in-person, open discussion prior to returning home to complete the second round online. As cited earlier, recent research suggests that incorporating a single discussion stage within a standard Delphi process generates improvements in response accuracy. This setup, therefore, allows us to make direct comparisons of degrees of participation and data quantity and quality between the two approaches.

2.1. Questionnaire design

The questionnaire on oyster ecosystem services was very involved and covered several topics, so it was important that it be designed in a way that eased the respondent into the elicitation exercise and began from a point of familiarity. Table 2 contains a summary outline of the major components of the questionnaire. The first section covered all basic and logistical information. First, respondents were asked to view and complete it on a full-screen monitor, rather than on a smart phone or tablet, because several of the questions and tables would be difficult to view on a small screen. They were then reassured of the anonymity of responses, reminded that the questionnaire was very demanding, encouraged to take breaks, and reminded that they had three weeks to complete it. The questionnaire also provided investigators' contact information. The oyster resources (as described earlier) and the ecosystem services to be covered by the questionnaire were then reviewed. Ecosystem services were *oyster density* (here, focused on the eastern oyster, *Crassostrea virginica*), *net nitrogen assimilation*, *reduced erosion*, *blue crab* (*Callinectes sapidus*) *density,* and *red drum* (*Sciaenops ocellatus*) *density* (see Table 1). The combinations of three resources and five ecosystem services means that responses were asked to respond regarding 15 resource-service combinations. The importance of complete responses, especially for oyster density levels, was then emphasized, but respondents were also reminded that they had the option to skip questions regarding resources and/or services with which they were unfamiliar. To mitigate concern about providing responses for

Table 2
Outline of questionnaire.

I. Introduction and logistical information
a. Provide screen requirements
b. Reminder of anonymity of responses
c. Reminder that questionnaire is demanding, encourage respondent to take breaks
d. Provide completion deadline
e. Provide contact information of the researchers.
f. Provide description of items to be covered by the questionnaire
g. Reminder about importance of complete responses
h. Discussion about providing responses for unfamiliar items, familiarity levels
II. Baseline (familiar) site from which all subsequent questions would pivot
a. Elicit baseline resource
b. Elicit baseline resource details
c. Elicit baseline resource familiarity level
III. Ecosystem service level data collection
a. Oyster density
i. Elicit service familiarity question
ii. Elicit preferred unit of measure
iii. Elicit absolute minimum and maximum bounds
iv. Elicit point estimate under baseline environmental scenario
v. Elicit point estimates for alternative salinity levels
vi. Elicit point estimates for alternative sedimentation levels
vii. Elicit point estimates for alternative nutrients-oxygen levels
viii. Review of all previous responses
ix. Respondent selects Resource 2
1. Elicit resource familiarity level
2. Same procedure as above
x. Respondent selects Resource 3
1. Elicit resource familiarity level
2. Same procedure as above
xi. Final review of all oyster density responses
b. Respondent selects Service 2; same procedure as above
c. Respondent selects Service 3; same procedure as above
d. Respondent selects Service 4; same procedure as above
e. Service 5; same procedure as above
IV. Elicit relative frequencies for 12 environmental scenarios
V. Closing screen: thank respondents and provide opportunity for comments

unfamiliar items, respondents were asked to provide their familiarity level with the item and reminded that familiarity level would be considered when responses were aggregated.

In the second section, respondents were asked to identify a specific site with an existing oyster resource on which they have worked and with which they were familiar. All subsequent questions would pivot off of this initial site and its prevailing environmental conditions. Respondents were asked to provide the name of the water body, the state in which the site is located, and the nearest city. Optionally, respondents could provide latitude/longitude and any other relevant information regarding the site. Next, respondents were asked to assess their own familiarity with the site (on a scale from 0 – not at all familiar – to 10 – very familiar). We acknowledge the limitations of a self-reported familiarity level, as more modest respondents may score themselves lower than less modest respondents with equal knowledge and experience. Next, environmental conditions typical at the site were elicited. Nutrients and oxygen conditions were elicited first, with quantitative definitions of terms provided, and respondents were asked to characterize the site's typical conditions (where typical was defined as "> 50 percent of the time") as either *eutrophic and hypoxic* or *mesotrophic and normoxic.* Sedimentation conditions were elicited next, with choices being *normal* or *high sedimentation.* Salinity levels were elicited last, with choices being *low (*< 10 ppt*), medium (10*–20 ppt*),* or *high salinity (*> 20 ppt*).* For salinity, respondents could indicate one, two, or all three salinity conditions as being typical at the site (see Table 1 for additional details). The combinations of two nutrient and oxygen levels, two sedimentation levels and three salinity levels means that respondents were asked to provide answers on 12 different environmental scenarios for each resource-service combination.

With the background information for the baseline site established, respondents were presented with specific questions about the initial resource-service combination. First, respondents were asked to identify the specific oyster resource present at the site. The familiarity question then asked respondents to indicate the level of familiarity with the chosen resource. For those responding with a level of familiarity less than three, an additional screen acknowledged their low level of familiarity, reminded them that familiarity levels would be considered, and then asked them to provide complete responses to the best of their ability.

The next section began the ecosystem services level data collection in earnest. The first set of questions focused on oyster density. Note that in the first round, oyster size was not specified. In response to the Discussion Group's recommendation, however, the Discussion Group considered oysters >25 mm only during Round 2. Familiarity with estimating oyster density was then elicited, then respondents were asked to provide their preferred unit of measure (number of oysters per square foot, number of oysters per square meter, or other specified by respondent). The next question sought to establish reasonable bounds for oyster density at the site by asking respondents to provide the absolute best (under the best environmental conditions expected at the site) and absolute worst (under the worst conditions expected at the site) oyster density they would reasonably expect to encounter at the site.

Next, the environmental conditions pivoted to the other two salinity conditions under the same nutrients-oxygen and sedimentation conditions, and embedded the previous response in the response cells, so that respondents could then input the oyster density estimates for the other salinity levels, as well as revise the previous response, if needed. This ability to revise as they worked continued throughout the questionnaire until all responses were solicited and only the final responses were those ultimately used. Conditions pivoted further to include the alternative sedimentation levels, with nutrients-oxygen levels still fixed. Finally, the alternative nutrients-oxygen conditions were presented. The complete table was then re-presented for review and revision, if necessary. At this point, the solicitation of expert responses for the twelve environmental scenarios for the first resource-service combination was complete.

Next, respondents were asked to select an alternative resource while keeping the ecosystem service fixed. Respondents were asked to consider a (potentially hypothetical) scenario where this alternative resource was present at the baseline site. The same line of questioning as described above was followed, except that the entire table of 12 environmental scenarios was presented at once, in the same arrangement as before. Note that for a different respondent that chose different environmental conditions, the table would be formatted differently, beginning with that respondent's typical conditions. Also presented were the previous responses provided for the first resource, so that the respondent could compare, and if necessary, revise, those as well. The same procedure then followed for the remaining (third) oyster resource, with both previous sets of responses shown for reference and any necessary revisions. A final screen presented all responses for all three resources for final review, and asked respondents to confirm that they were aware that this was the last opportunity to revise the values for that particular ecosystem service.

The remainder of the questionnaire focused on the remaining ecosystem services: net nitrogen assimilation, reduced erosion, blue crab density, and red drum density. The respondent was asked to choose one of the remaining services; we recommended they begin with the one with which they were most familiar. As before, familiarity level was also collected, and those with low levels of familiarity were given the option to skip this line of questioning. As before, upper and lower bounds were first elicited, then the 12 environmental scenarios. For each question, the corresponding values for the first ecosystem service (oyster density) provided in the previous section were presented for reference (but, at this point, not for revision). The same line of questioning was then followed for the next service chosen, then the remaining one. At this point, the solicitation of expert responses for the twelve environmental scenarios for all 15 resource-service combinations is complete. The final question elicited the relative frequencies of observing the 12 environmental scenarios at the respondent's site. Thus, the complete set of responses provided the probability of observing each environmental scenario at the site, as well as the estimated levels of each service under each resource conditional on each environmental scenario being observed. The closing screen thanked respondents and provided an opportunity to provide any comments.

2.2. Questionnaire testing, revision, sample, and administration

The questionnaire was designed using the Qualtrics online survey platform, tested, and revised over a 10-month period, from October 2017 to July 2018. Expert panelists, including external reviewers, were recruited in June 2018. Four external reviewers reviewed the draft questionnaire during June and July, and suggested revisions were incorporated.

Our sample was constructed as follows. We surveyed the literature to identify experts with a relevant publication record. In addition, we created a broad list of any additional individuals with interest in the topic and a strong likelihood of familiarity with the systems being studied. Our list was stratified by state, gender, and sector (e.g., academia, natural resource management, commercial industry, etc.). We strove to be as inclusive as possible in our search, and believe it is representative of the two regions' experts. In terms of familiarity, most respondents often had 2-3 sites with which they were most familiar, with some experts having familiarity with a broader range of sites. We did not specifically attempt to include generalists, but believe that several of the included experts could be characterized that way.

The final Round 1 questionnaire, which was identical for the Isolated and Discussion groups, was administered to 39 expert panelists during July and August 2018, with periodic reminders sent to non-respondents. One panelist opted out shortly thereafter. Completion times ranged from 1 to 18 days. Round 1 response summaries were sent to all panelists on August 24. The summary included tables of summary statistics (N, mean, minimum, maximum), as well as box-and-whisker plots of responses, for each resource-service combination. The Discussion Group's in-person panel was held August 31, 2018, at the Dauphin Island Sea Lab, Alabama. That same day, the questionnaire was re-sent to Isolated Group respondents to make any revisions based on the response summaries, and was in the field from September to November, with reminders sent in-between. The questionnaire remained unchanged for the Isolated Group. For Discussion Group respondents, the questionnaire was revised (revisions described below) based on the in-person panel, and this revised questionnaire was sent back to them to complete, remaining in the field from September to November, with periodic reminders sent to non-respondents. All responses were received by the end of November 2018. Data review and processing began shortly thereafter. Problematic responses were identified, and follow-up email correspondence was sent to individual panelists as needed. All correspondence with panelists was concluded by the end of March 2019 and the dataset was finalized.

Seven clarifying revisions came out of the in-person discussion. The first added a question where, for *restored*, the respondent indicates substrate type, whether harvest is allowed, and whether they are subtidal or intertidal. The second defined hypoxic conditions as "sub-lethal effects on adults over a one-week timespan". The third specified that only oysters >25 mm in size be considered. The fourth separated blue crab responses into two size classes: 0–30 mm (small) and >30 mm (large juveniles and adults). The fifth separated red drum into three size classes: 25–150 mm (small), 150–350 mm (medium), and >350 mm (large). The sixth defined the relevant nitrogen metric as "net nitrogen assimilation into oyster tissue and shell". The seventh added a question regarding erosion, that elicited a baseline erosion rate, and revised scenario responses to be observed rates rather than net changes.

2.3. Literature as potential test of external validity

Prior to the design of the expert questionnaire, a comprehensive literature review was conducted for the ecosystem services provided by *traditional*, *restored*, and *off-bottom* oyster resources. Over 200 publications and resources were sorted through and categorized. The initial plan was to use the results of a literature meta-analysis as a test of external validity of the expert elicitation estimates. The literature review was limited to quantitative studies that would allow for comparison of units across papers. The majority of the literature found was focused on *restored*, followed by *traditional*, then *off-bottom*. Regarding ecosystem services, the largest number of papers focused on habitat, followed by oyster harvest or abundance. Papers referencing water quality and erosion protection were relatively scarce. Metrics and reporting of statistics also varied widely. For example, although we found roughly 125 quantitative studies about water quality, only 50 of those used metrics that could be compared across at least two other publications, and reported basic summary statistics (number of observations, mean, and standard deviation). In the end, we concluded that a meta-analysis for the purpose of comparing to our expert data was infeasible, so we merely report the results of the literature search and provide a summary of findings in Appendix A.

3. Results

3.1. Sample overview

The final sample consisted of 38 individuals, with 19 from the U.S. Atlantic Coast and 19 from the U.S. Gulf Coast. Of these, 22 were assigned to the Isolated Group, and 16 were assigned to the Discussion Group. Thirty were academics, 8 were from federal or state government agencies (working as resource managers), and 4 were from non-governmental organizations or industry.

3.2. Estimates by round, elicitation method, and region

We first provide a discussion of results related to our second research question *Do oyster resource estimates vary based on region and elicitation method?* Table 3 reports means and standard deviations of estimated ecosystem service levels over all environmental scenarios by oyster resource, round, and subsample, that is, by elicitation type (Isolated or Discussion) and region (Gulf or Atlantic). Also reported are the results of two-sided t-tests of mean differences between rounds, elicitation subsamples, and region subsamples, respectively. (Note that unpaired t-tests were consistent across assumptions of equal and unequal variance in all but four cases. In these cases, where standard deviations were found to be significantly different, we report the *t*-test result for unequal variance, and vice-versa.) The results of F-tests of standard deviation differences between rounds are also reported. Because respondents could opt out of resources and services for which they were less familiar, individual estimates are based on different numbers of observations. These differences were most pronounced across ecosystem services, although we generally observed slightly lower numbers of responses for *off-bottom* relative to the other resources. Out of 38 respondents, oyster density averaged 36 respondents per scenario; blue crab density averaged 28; reduced erosion averaged 24; net nitrogen assimilated averaged 20, and red drum density averaged 10. The full set of summary statistics for these data are reported in Appendix B. Note that the means reported in this section should not be interpreted as the levels expected to be observed because environmental scenarios are given equal weight; they do not account for relative probabilities of each environmental scenario occurring. These means are reported only for the purpose of comparing responses across rounds and subsamples. The subsequent section will report scenario-specific levels.

We find only two instances (both for red drum) of significant (95 percent level) differences in means between rounds for the Isolated subsample, and no differences in standard deviation. For the Discussion subsample, however, we find substantial differences. We find eleven instances of significant mean differences and twelve instances of significant standard deviation differences between rounds. In all but two of these instances, standard deviations decrease in Round 2.

Comparing the elicitation subsamples, in Round 1 we find 9 instances of significant mean differences and 12 instances of standard deviation differences. In the 3 instances where the mean was not different but the standard deviation was, the Discussion subsample had the higher standard deviation in 2 of them. In Round 2, the number of mean differences drops to 6, but the number of standard deviation differences remains at 12. In the 8 instances where the mean was not different but the standard deviation was, the standard deviation for the Discussion subsample was lower in all but 2. Taken together, these results indicate that the in-person panel discussion appears to induce a greater degree of engagement and subsequent revisions in initial responses, whereas the isolated response approach leads to no real revisions. Additionally, the discussion approach appears to lead to reductions in variation of estimates relative to the isolated approach.

We also compare estimates across regional subsamples, as it is possible that ecosystem service levels may differ by region. Note that elicitation subsamples are pooled in this case. Note also that because Round 2 responses represent the final responses, we leave behind Round 1 responses from this point forward and focus on Round 2. We find 6 instances of significant mean differences between region subsamples. Mean net nitrogen assimilated for *off-bottom* is significantly lower in the Atlantic subsample, as is reduced erosion for all resources, and blue crab density for *off-bottom*; red drum density is higher for *traditional*. No significant differences are found for oyster density for any resource.

The final comparison made in this section is across. Results indicate general agreement among 3 of the 4 subsamples that each resource has a distinct oyster density level, with *off-bottom* having the highest oyster density level, followed by *restored*, then *traditional*. The Discussion subsample also indicates that *off-bottom* has a significantly higher level, but with no significant difference between *traditional* and *off-bottom*.

For net nitrogen assimilated, results indicate that *traditional* levels are consistently different than *restored* levels, with *off-bottom* levels falling generally in-between, although the relative magnitudes are not

Table 3
Frequency (N), means (μ), and standard deviations (σ) of service levels by resource, round, and subsample.

Service	Resource	Round 1		Round 2			
		Isolated	Discussion	Isolated	Discussion	Gulf	Atlantic
Oyster Density	Traditional	N = 240	192	240	192	212	220
(#/m^2)		μ = 83.89	112.00	83.87[d]	74.62[a,b]	75.41[d]	83.95[d]
		σ = (250.30)	(304.20)[b]	(250.80)	(232.60)[a]	(263.20)	(221.60)
	Off-bottom	240	192	240	192	220	212
		230.00	188.20	229.60[d]	220[d]	243.60[d]	206.40[d]
		(282.50)	(276.60)	(282.40)	(259.50)	(300.90)	(238.00)
	Restored	242	192	242	192	214	220
		157.30	172.30	170.40[d]	73.79[a,b]	123.90[d]	131.30[d]
		(528.90)	(464.60)	(542.40)	(232.20)[a,b]	(528.80)	(321.50)
Net N Assimilated	Traditional	141	132	141	108	84	165
(g N/m^2)		89.19	200.10[b]	88.99[d]	81.58[a]	66.38[d]	95.64
		(157.00)	(611.30)[b]	(157.10)	(282.40)[a,b]	(151.70)	(247.40)
	Off-bottom	128	132	128	108	84	152
		79.03	140.60[b]	76.49	56.69[a]	91.54	54.11[c]
		(165.10)	(293.50)[b]	(164.90)	(87.32)[a,b]	(200.30)	(76.67)
	Restored	142	132	142	108	84	166
		105.50	262.80	105.40	41.31[b,d]	95.73	68.62[d]
		(205.60)	(945.60)[b]	(205.70)	(98.61)[a,b]	(234.10)	(127.10)
Reduced Erosion	Traditional	192	132	192	108	144	156
(m shoreline/m^2)		0.32	0.41	0.32	0.31	0.53	0.11[c]
		(1.80)	(1.83)	(1.80)	(0.61)[a,b]	(2.09)	(0.34)
	Off-bottom	179	108	179	96	131	144
		0.23	0.74[b]	0.23	0.33[a]	0.42	0.12[c]
		(1.00)	(1.58)[b]	(1.00)	(0.83)[a,b]	(1.29)	(0.39)
	Restored	195	132	195	108	148	155
		1.60[d]	0.47[b]	1.60	0.53[b,d]	2.22[d]	0.26[c,d]
		(7.08)	(2.02)[b]	(7.08)	(0.64)[a,b]	(8.04)	(0.44)
Blue Crab Density	Traditional	218	168	218	132	165	185
(#/m^2)		6.16	2.66[b]	6.20	2.85[a,b]	5.21	4.69
		(14.56)	(4.04)[b]	(14.56)	(3.89)[b]	(10.98)	(12.57)
	Off-bottom	188	140	188	118	152	154
		4.09	3.01	4.18	2.94[a]	6.07	1.37[c]
		(10.04)	(4.79)[b]	(10.08)	(4.66)[b]	(11.35)	(1.80)
	Restored	226	168	226	132	172	186
		5.74	3.18[b]	5.84	3.11[a,b]	5.08	4.61
		(13.06)	(5.39)[b]	(13.07)	(4.20)[a,b]	(8.80)	(12.33)
Red Drum Density	Traditional	48	132	48	84	84	48
(#/m^2)		0.88	0.29[b]	0.75[a]	0.58[a]	0.49	0.90[c,d]
		(1.11)	(0.60)[b]	(1.15)	(1.12)[a]	(0.92)	(1.40)
	Off-bottom	36	108	36	72	72	36
		0.86	0.17[b]	0.78	0.84[a]	0.87	0.73[d]
		(0.92)	(0.43)[b]	(0.95)	(1.35)[a,b]	(1.33)	(1.00)
	Restored	48	132	48	84	84	48
		0.93	0.25[b]	0.79[a]	0.46[a,d]	0.51[d]	0.71[d]
		(1.11)	(0.53)[b]	(1.14)	(0.89)[a,b]	(0.87)	(1.18)

[a] Round 2 mean (standard deviation) significantly different (95% level) than Round 1 mean (standard deviation) based on two-sided *t*-test (F-test), within elicitation subsample.
[b] Discussion subsample mean (standard deviation) significantly different than Isolated subsample mean (standard deviation) based on two-sided *t*-test (F-test), within round.
[c] Atlantic subsample mean significantly different than Gulf subsample mean based on two-sided *t*-test (Round 2 only).
[d] Resource mean significantly different from each of the other resource means based on pairwise two-sided t-tests, within service and subsample (Round 2 only).

consistent across subsamples. Note that summary statistics are based on all observations, whereas we used paired t-tests, where applicable, some of which are based on a smaller number of observations. Consequently, the reported test results may appear inconsistent with the reported summary statistics in a few cases. Reduced erosion levels are consistently significantly greater for *restored*, whereas levels for *traditional* and *off-bottom* are not significantly different. We find some significant differences for blue crab density across resources. Red drum density levels are more mixed. The Isolated subsample indicates no significant differences, whereas the Discussion and Gulf subsamples indicates relatively higher levels for *off-bottom*, intermediate levels for *traditional*, and the lowest levels for *restored*. The Atlantic subsample, however, indicates distinct levels across resources, with the highest levels for *traditional*, followed by *off-bottom*, then *restored*.

To provide some additional guidance to users of our data, we also investigated whether the data exhibited any systematic non-response patterns, insofar as non-response may be influencing our estimates. We estimated logistic regressions for each service, with a binary indicator for response (versus non-response) dependent variable and round, elicitation type, and region as binary explanatory variables, using cluster-robust standard errors (clustered on respondent). We found no significant non-response patterns for oyster density, reduced erosion, or blue crab density. However, we found that Gulf respondents were significantly (p-value < 0.05) less likely to provide net nitrogen assimilated responses. This implies that the net nitrogen assimilated estimates are potentially biased toward the Atlantic region. Concerning elicitation type, we found that Isolated respondents were significantly less likely to provide red drum density responses. The implications of this latter finding is not clear, but possibly reflects unobserved idiosyncrasies of individual respondents that transcends our experimental controls. We can only emphasize that readers use caution when interpreting and using these estimates.

In summary, we find that estimates may indeed differ depending on elicitation method and region, and we make no judgment on which set of results are superior. We simply report them as is so that other researchers and managers can evaluate and utilize them when evaluating potential ecosystem services from different oyster resources. As a reminder, we also provide Appendix B, which contains full summary statistics for every combination of environmental scenario, ecosystem service, oyster resource, and elicitation subsample.

3.3. Estimates by environmental scenario

To give the reader a more in-depth understanding of the differences discussed above, Figs. 1–3 plot median, mean, and familiarity-weighted mean oyster density levels for each resource under each of the twelve environmental scenarios by subsample, compared to the pooled ("All") sample. Familiarity weights were calculated as follows: for each respondent, each resource familiarity score was multiplied by each service familiarity score for every resource-service combination. Then, the familiarity score products were summed over those respondents that provided a response for that particular round-resource-service-scenario combination. Then, the ratio of each familiarity score product and the sum of familiarity score products were calculated to yield the familiarity weights. The weighted means were then calculated using these familiarity weights. Median levels tend to be lower than mean levels across the board, and we find more consistency across subsamples for *off-bottom*. For *traditional* and *restored*, means across subsamples tend to differ most for the eutrophic-hypoxic/normal sedimentation/medium and high salinity scenarios (ENM and ENH), and for *restored*, for the mesotrophic-normoxic/normal sedimentation/medium and high salinity scenarios (MNM and MNH). Median differences for *traditional* and *restored* tend to occur for the mesotrophic-normoxic/normal sedimentation/low, medium, and high salinity scenarios. Appendix C contains similar figures (Figs. C1-C12) for net nitrogen assimilated, reduced erosion, blue crab density, and red drum density.

3.4. Estimates by resource

The other central question this work sought to answer was *Do ecosystem services differ among oyster resources?* Acknowledging the aforementioned differences found in our estimates, to facilitate discussion, from this point forward, we focus on the pooled ("All") estimates. Figs. 4–8 plot median, mean, and familiarity-weighted mean ecosystem service levels under each of the twelve environmental scenarios by resource. Fig. 4 indicates that oyster density is consistently higher for *off-bottom*. Median estimates indicate larger differences under mesotrophic-normoxic/normal sedimentation/medium and high salinity scenarios (MNM and MNH) than do mean estimates, whereas mean estimates indicate larger differences under mesotrophic-normoxic/normal and high sedimentation/low salinity scenarios (MNL and MHL).

Fig. 5 indicates more muted differences for net nitrogen assimilated, with relatively larger differences observed under mesotrophic-normoxic scenarios, although these differences are further muted based on familiarity-weighted mean estimates. Fig. 6 indicates consistently higher erosion reductions for *restored*. Median estimates indicate that the differences across resources are large under all except one scenario (EHH), whereas mean differences indicate that differences are scenario-specific.

Fig. 7 indicates that relative blue crab density is more mixed and scenario-specific. Each resource is shown with the highest density under at least one scenario/estimate type. Fig. 8 indicates that relative red drum density is also mixed, but *off-bottom* tends to have higher density levels under eutrophic-hypoxic scenarios. Under mesotrophic-normoxic scenarios, median estimates tend to favor *restored*, whereas mean estimates are mixed.

3.5. Summary metrics of relative performance by resource

In this section we wish to assess overall relative performance across oyster resources. We use two summary measures, percent deviation from mean ecosystem service level and coefficient of variation. Percent deviation from the mean is defined as the ratio of the resource-specific mean service level across scenarios and the mean service level across resources, minus one. Positive values indicate above-average performance and negative values indicate below-average performance. For example, one particular respondent's mean oyster densities over all scenarios for the three resources were 92.08 (*traditional)*, 337.50 (*off-bottom)*, and 84.17 (*restored*). So, this respondent's percent deviation from the mean for *traditional* is {92.08/[(92.08 + 337.50 + 84.17)/3]} - 1 = −0.46, indicating a 47 percent below-average performance.

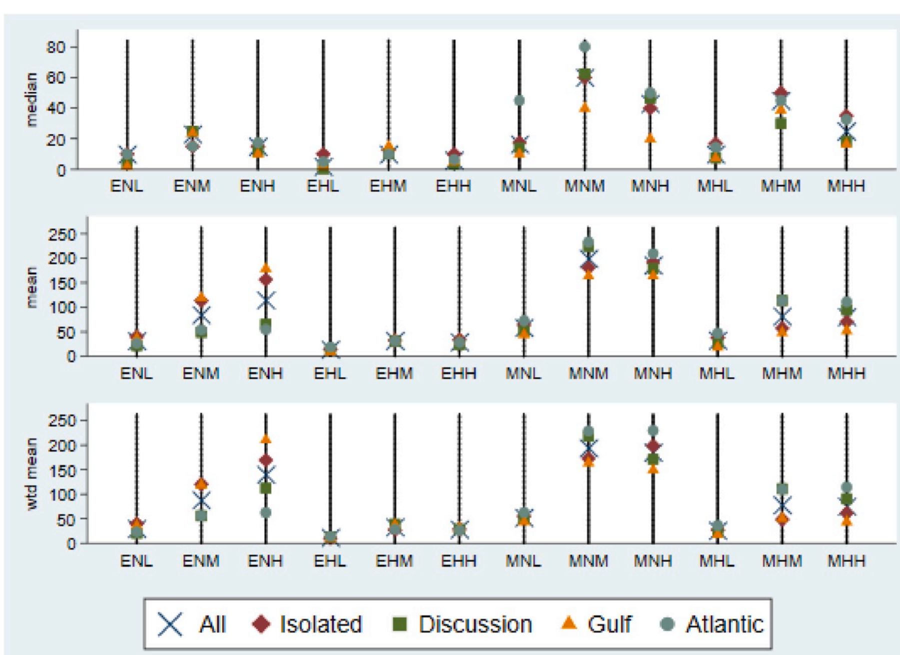

Fig. 1. Comparison of oyster density (#/m^2) by subsample, *traditional* oyster resource, over aggregation method (panels) and environmental scenario (columns).

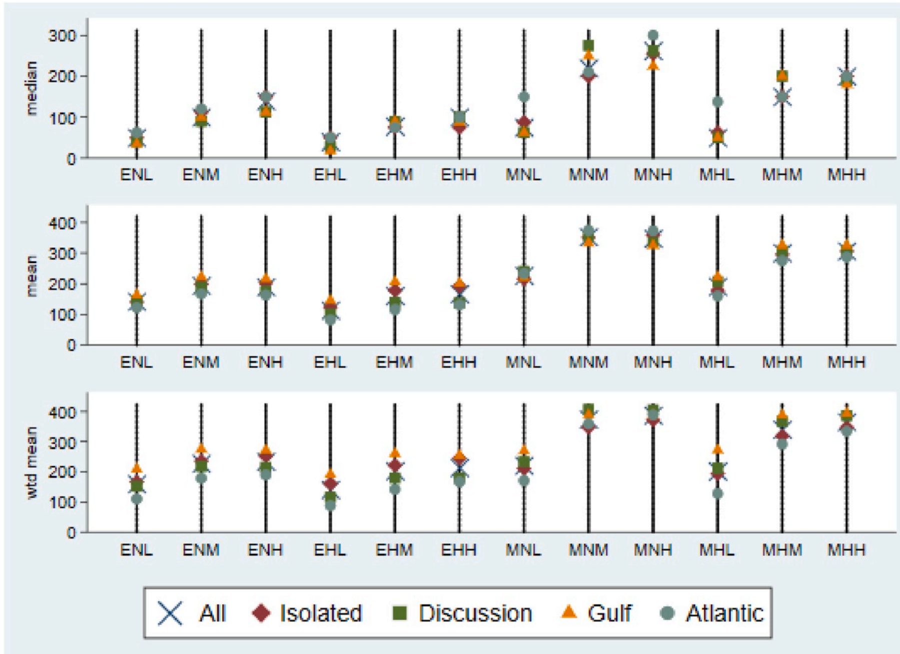

Fig. 2. Comparison of oyster density (#/m^2) by subsample, *off-bottom* oyster resource, over aggregation method (panels) and environmental scenario (columns).

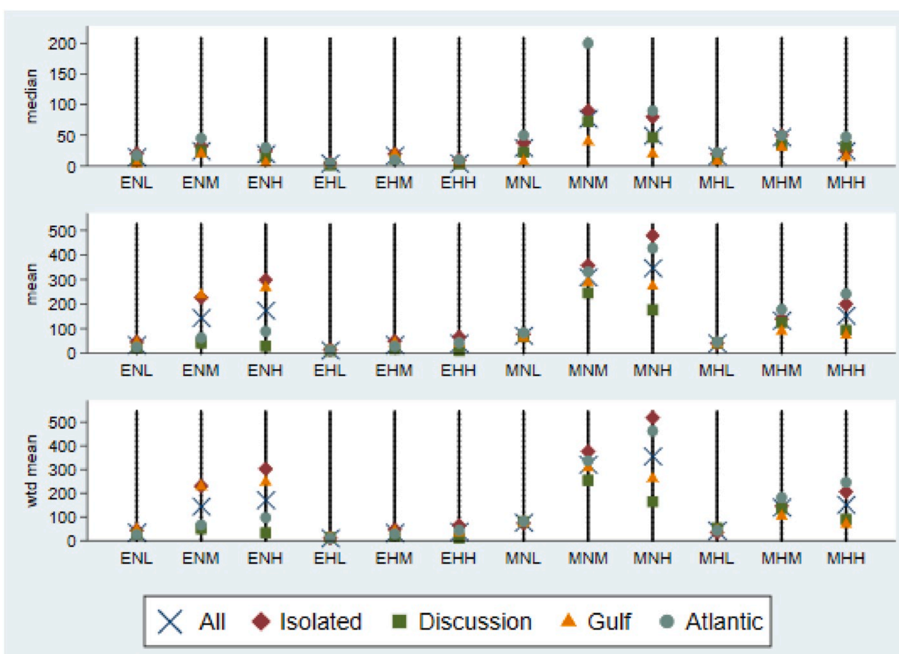

Fig. 3. Comparison of oyster density (#/m^2) by subsample, *restored* oyster resource, over aggregation method (panels) and environmental scenario (columns).

Coefficient of variation for a given resource is defined as the ratio of standard deviation and mean. Again, this particular respondent's standard deviation for *traditional* was 146.68, so coefficient of variation is (146.68/92.08) = 1.59.

Now, there are two sources of variation in our data. The first is variation across scenarios, *within subject.* This represents the "true" variation in service levels (as respondents perceive them) due to changes in environmental conditions. The second source of variation is *between subjects.* This represents the uncertainty among respondents regarding the true values. In this section, we wish to control for the latter and examine the former. To do so, we calculate percent deviation from the mean and coefficient of variation *within subjects* first. Then, we calculate the means and standard deviations of these performance measures. In this way, the mean values reflect the *within subject* variation, that is, the average respondent's estimates of relative performance. We then construct confidence intervals around these means, and the confidence intervals reflect the *between subject* variation, that is, the degree to which respondents differed in their responses.

Percent deviation from the mean results are reported in the top panel of Fig. 9. Here we pool observations because pooled results are fairly consistent with those of individual region and elicitation subsamples. The reader is directed to Appendix C, Figs. C13-C16, for the results by

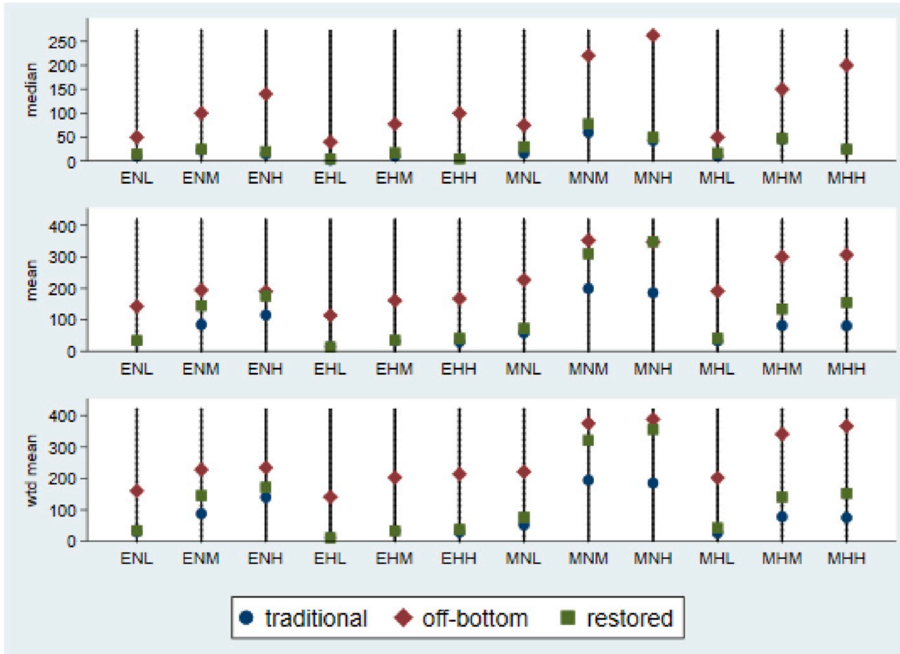

Fig. 4. Comparison of oyster density (#/m^2) by oyster resource, over aggregation method (panels) and environmental scenario (columns).

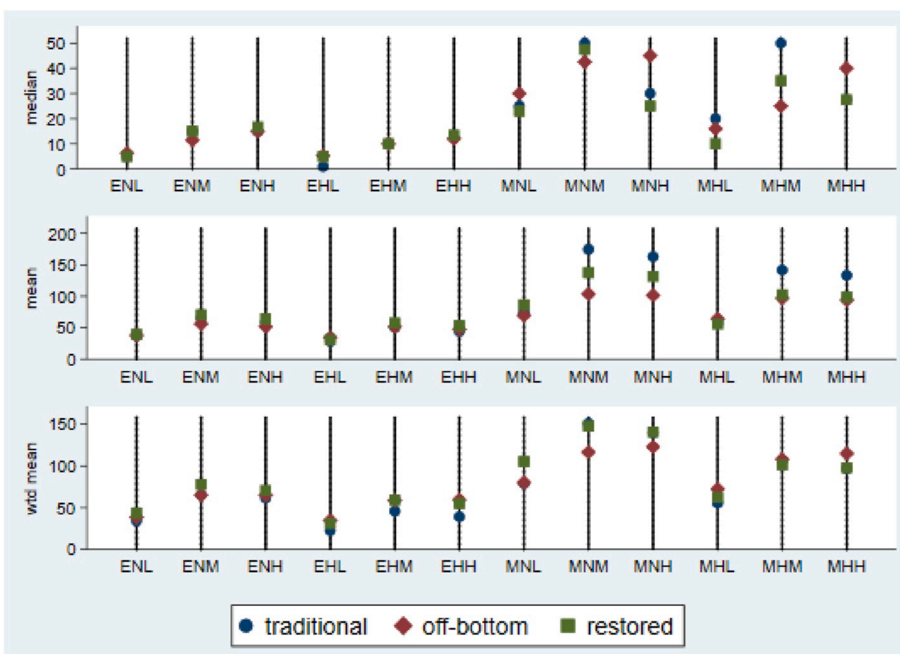

Fig. 5. Comparison of net nitrogen assimilated (g N/m^2) by oyster resource, over aggregation method (panels) and environmental scenario (columns).

subsample. Note that we exclude respondents that did not provide responses for all 12 scenarios for a given resource-service combination. The origin of the vertical axis in Fig. 9 represents the mean level of ecosystem service delivery. Thus, our data indicate that *traditional* and *restored* deliver, on average, a 42 percent and 34 percent lower level of oyster density, respectively, relative to the three-resource average, whereas *off-bottom* delivers, on average, a 73 percent higher level of oyster density. Furthermore, the confidence-interval (CI) bars show that the difference between *off-bottom* and the other two resources is significant, that is, that.

respondents systematically estimated *off-bottom* oyster density at higher levels than the other two resources. For the remaining services, although there are some visible differences in performance, the CI bars show substantial overlap, indicating that these differences are not significant. Results using familiarity-weighted means are very similar, so we do not discuss them here, but report them in Appendix C, Fig. C17.

Coefficients of variation are reported in the bottom panel of Fig. 9. Results indicate that, on average, coefficient of variation was lower for *off-bottom* relative to *traditional* and *restored* across all services. In other words, on average, there was less variation in *off-bottom* service levels across environmental scenarios. Additionally, the CI bars indicate that the oyster density difference is significant across resources, that is, that

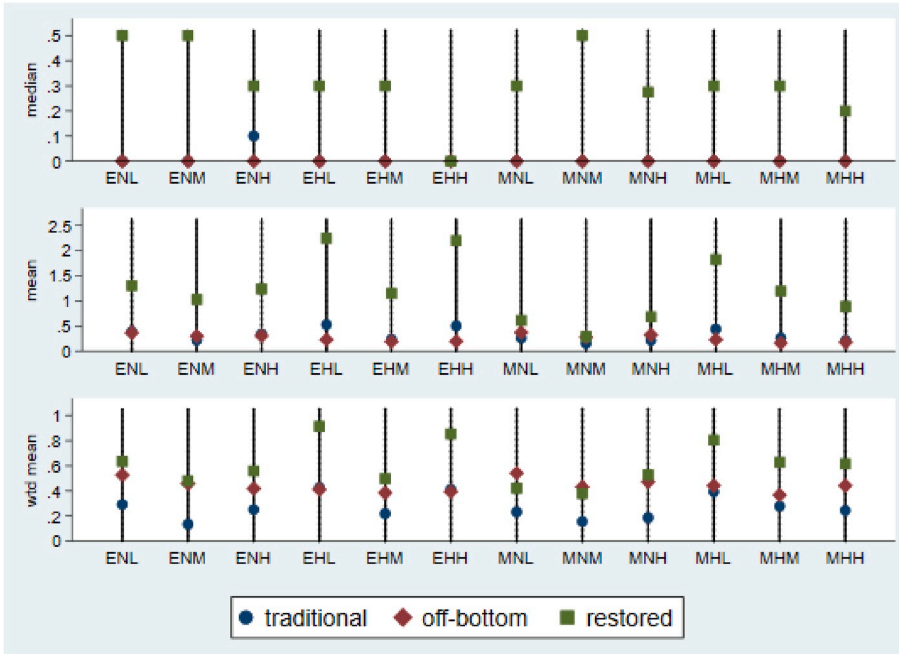

Fig. 6. Comparison of reduced erosion (m shoreline/m^2) by oyster resource, over aggregation method (panels) and environmental scenario (columns).

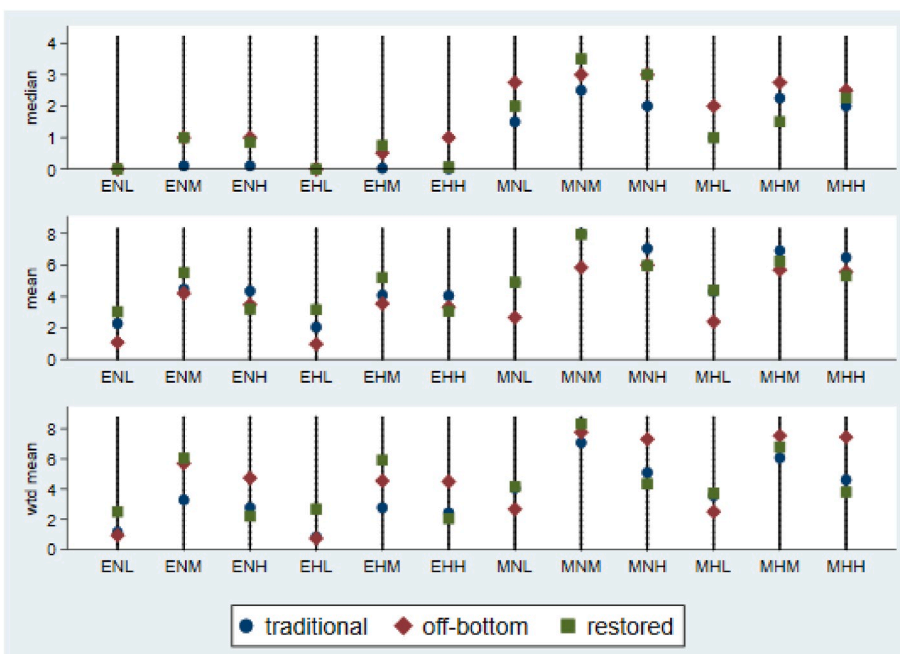

Fig. 7. Comparison of blue crab density (#/m^2) by oyster resource, over aggregation method (panels) and environmental scenario (columns). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

respondents systematically estimated lower variation in oyster density for *off-bottom* relative to *traditional* and *restored*. CI bars for the remaining services show substantial overlap, indicating that none of the other mean differences in coefficient of variation are significant.

4. Discussion

Results from our elicitation treatments indicate that elicitation method can indeed affect the distribution of final responses. However, we note that even during Round 1, in which there were no differences in elicitation procedures, we found significant mean differences for 8 of the 15 estimates elicited. Fewer differences are observed when comparing elicitation types directly, and more when comparing how responses changed *within* each elicitation treatment going from Round 1 to Round 2. We find only two instances of significant differences in means between rounds for the Isolated subsample, and no differences in standard deviation. For the Discussion subsample, however, we find eleven instances of significant mean differences and twelve instances of significant standard deviation differences between rounds. Thus, we find that the Discussion treatment results in substantially higher revision rates

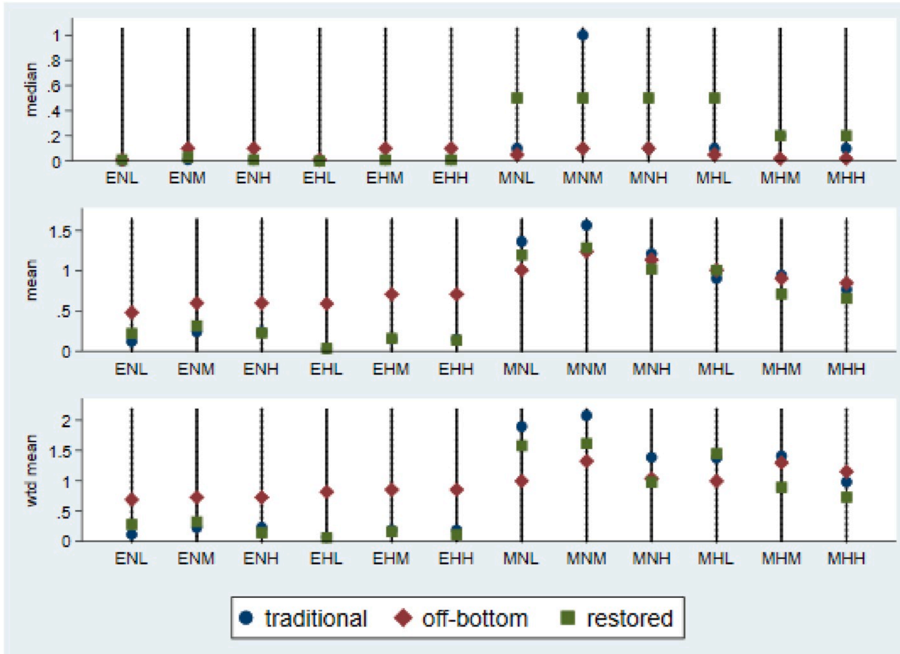

Fig. 8. Comparison of red drum crab density (#/m^2) by oyster resource, over aggregation method (panels) and environmental scenario (columns). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 9. Top: Percent deviation from mean ecosystem service level (top) and coefficient of variation (bottom) by oyster resource as a measure of relative level of ecosystem services.

between rounds.

We caution, however, that these results do not necessarily imply that the Discussion treatment was the superior approach. Our group discussion, like any group discussion, featured individuals that spoke frequently and confidently, and others that spoke little and/or more tentatively. So it is possible that a small number of researchers' opinions had undue influence on others such that their final estimates were inconsistent with their own personal knowledge and experience. It should be noted, however, that respondents still provided final responses individually and privately, so that peer pressure should have not played a major role.

We also found some significant higher service estimates among Gulf respondents relative to Atlantic respondents: mean net nitrogen assimilation for *off-bottom* was 69 percent higher among Gulf respondents; reduced erosion was 250–750 percent higher, depending on resource; and blue crab density was 343 percent higher for *off-bottom*. The remaining significant difference found red drum density 84 percent higher for *traditional* among Atlantic respondents. These results are intriguing and suggest that further work is necessary to determine if these results reflect true differences in ecosystem services between regions or simply differences across expert respondents' perceptions. In the effort to quantify ecosystem services, it will be essential to understand the impact of geographic region, if any.

Our summary performance metrics tell a slightly different story. First, as a matter of convenience, we present these results aggregated across subsamples because it would be too much to present these for each individual subsample (but the reader can examine them in Appendix C). Furthermore, although there are indeed some differences in particular cases across subsamples, the general finding is the same: we find significant differences for oyster density only, with *off-bottom* outperforming both *traditional* and *restored* in terms of both higher mean density and lower variability. For the remaining services, however, we find substantial overlap in confidence intervals, indicating perhaps that the true differences are not large, that there is substantial disagreement among our experts, or that we simply do not know what the true values are.

Given the potential differences between the regional responses as well as the elicitation methods, a manager might ask how these data should be utilized. First, as we suggest above, the regional differences observed warrant further exploration to determine if those differences reflect real differences in those natural systems or if those differences are due to differences in respondents' perceptions. Second, we urge caution when interpreting the data from either elicitation method and suggest that managers use these data as current best estimates. These estimates of ecosystem services provide a valuable starting point, either as a point of comparison for locally collected field data or as a reasonably justified assumption in the absence of such data. Given the cost of the field studies required to assess this range of ecosystem services, the data generated here provide value. Second, despite some differences between elicitation methods in terms of specifics, the general pattern is consistent that valuable ecosystem services are provided by each of the oyster resources evaluated. We also note in passing that these performance metrics should not be construed to indicate *economic* performance. Economic performance depends not only on quantities, but on market prices, ecosystem service values, and production costs. Future work would need to take all of these other factors into account to speak to economic performance of oyster resources.

Our data indicate that, among the oyster resources studied here, *off-bottom* is the one for which knowledge and experience is relatively scarce, and among services, reduced erosion and red drum density are the ones for which knowledge and experience is either relatively scarce or lacks consensus. Response rates for *off-bottom* tended to be lower than those of *traditional* and *restored*, although these differences were small and not universal. Responses rates for red drum density, however, were substantially lower, regardless of subsample, indicating a relative lack of knowledge and experience with this service. And the very wide confidence intervals on reduced erosion performance metrics indicate vast differences of opinion across responses. If these findings are representative of the larger population of oyster scientists and practitioners, then these are particular areas of need (and opportunity) for researchers in the days ahead. More generally, however, we note that standard deviations of service level estimates were larger than the means *across the board*, indicating that much work remains regarding understanding ecosystem service delivery by oysters.

5. Conclusions

Acknowledging that our study is not a complete substitute for much-needed empirical data – data, we should add, that would be very costly and would require many years of field work, lab work, and analysis – we contend it still presents the most comprehensive set of ecosystem service estimates for diverse (harvested and non-harvested) oyster resources to date. The gap in information regarding ecosystem services provided by harvested oyster resources, and how it compares with those provided by non-harvested resources, is particularly acute given the demand for this kind of information by scientists, managers, and regulators, especially when faced with decisions about multiple oyster resources and potential user conflicts (Baggett et al., 2015; Sutton-Grier and Bamford, 2015). We believe that our efforts and estimates here will bridge this information gap and will greatly enlighten the ongoing debate regarding the relative merits of different oyster production systems and the challenges associated with managing multiple oyster resources to achieve multiple objectives. Furthermore, the data highlight where empirical data may be needed most.

We also acknowledge some of the key weaknesses of our approach. First, we have attempted to cover a broad range of environmental conditions representative of the spatial and temporal variability in such conditions encountered in the Gulf and Atlantic coasts. Although our analysis could therefore be interpreted as capturing localized conditions and temporal variability, we acknowledge that our scenarios are mere proxies for the true spatial and temporal variability. A more in-depth analysis of how explicit types of temporal variability (e.g. monthly vs. seasonal vs. interannual) could affect such estimates is needed. Second, the analyses presented in this paper do not allow us to estimate possible dependencies among the ecosystem services targeted, as well as how to scale up our results to larger spatial extents. For example, enhanced shoreline protection can increase blue crab density through new growth/expansion of adjacent marsh (McDonald et al., 2016). Finally, other scaling non-linearities could affect our conclusions and utilization by managers. For instance, our results could get amplified under synergistic interactions, or else plateau under functional saturation, as reefs grow bigger and bigger (Barbier et al., 2008). More research is needed to shed light on these important questions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We wish to thank our expert panelists for participating and contributing their knowledge, experience, and responses. Our respondents were Christine Angelini, Donna Bilkovic, Seth Blitch, Suzanne Bricker, Ruth Carmichael, Dave Eggleston, Carl Ferraro, Stuart Findlay, Jon Grabowski, Boze Hancock, Ken Heck, Jason Herrmann, Lisa Kellogg, Kelly Kibler, Megan La Peyre, Dale Leavitt, Rom Lipcius, Mark Luckenbach, Giovanna McClenachan, Scott Milroy, Jenny Paterno Shinn, Mark Peterson, Jennifer Pollack, Sean Powers, Steven Scyphers, Rochelle Seitz, John Supan, Seth Theuerkauf, Linda Walters, Bret Webb, Chuck Wilson, Chester Zarnoch, and six respondents who wish to remain anonymous. Our expert questionnaire reviewers were David Bushek, Loren Coen, Julie Rose, and Leslie Sturmer. Funding for this article comes from the National Oceanic and Atmospheric Administration's RESTORE Science Program under award NA17NOS4510090 to Mississippi State University. This work was also supported by the National Institute of Food and Agriculture and the Mississippi Agricultural and Forestry Experiment Station via Multistate Project W-4133 "Costs and Benefits of Natural Resources on Public and Private Lands: Management, Economic Valuation, and Integrated Decision-Making", (Hatch Project MIS-033140).

Appendices A–C. Supplementary data

Supplementary documentation can be found online at https://doi.org/10.1016/j.jenvman.2020.110676.

References

Alleway, H.K., Gillies, C.L., Bishop, M.J., Gentry, R.R., Theuerkauf, S.J., Jones, R., 2019. The ecosystem services of marine aquaculture: valuing Benefits to people and nature. Bioscience 69 (1), 59–68.

Baggett, L.P., Powers, S.P., Brumbaugh, R.D., Coen, L.D., DeAngelis, B.M., Greene, J.K., Hancock, B.T., Morlock, S.M., Allen, B.L., Breitburg, D.L., Bushek, D., Grabowski, J. H., Grizzle, R.E., Grosholz, E.D., LaPeyre, M.K., Luckenbach, M.W., McGraw, K.A., Piehler, M.F., Westby, S.R., zu Ermgassen, P.S., 2015. Guidelines for evaluating performance of oyster habitat restoration. Restor. Ecol. 23, 737–745.

Barbier, E.B., Koch, E.W., Silliman, B.R., Hacker, S.D., Wolanski, E., Primavera, J., Granek, Ef, S, Polasky, Aswani, S., Cramer, L.A., Stoms, D.M., Kennedy, C.J., Bael, D., Kappel, C.V., Perillo, G.M.E., Reed, D.J., 2008. Coastal ecosystem-based management with nonlinear ecological functions and values. Science 319 (5861), 321–323.

Curtis, I.A., 2004. Valuing ecosystem goods and services: a new approach using a surrogate market and the combination of a multiple criteria analysis and a Delphi panel to assign weights to the attributes. Ecol. Econ. 50, 163–194.

Dalkey, N.C., 1967. Delphi. Rand Corporation, Santa Monica, CA. Rand Paper P-3704, October.

Dalkey, N.C., 1969. The Delphi Method: an Experimental Study of Group Opinion. June. Rand Corporation, Santa Monica, CA. Rand Paper RM5888-PR.

Dalkey, N., Helmer, O., 1963. An experimental application of the Delphi method to the use of experts. Manag. Sci. 9 (3), 458–467.

Donlan, C.J., Wingfield, D.K., Crowder, L.B., Wilcox, C., 2010. Using expert opinion surveys to rank threats to endangered species: a case study with sea turtles. Conserv. Biol. 24 (6), 1586–1595.

Evans, C., 1997. The use of consensus methods and expert panels in pharmacoeconomic studies practical applications and methodological shortcomings. Pharmacoeconomics 12 (2 Pt 1), 121–129.

Fodrie, F.J., Rodriguez, A.B., Gittman, R.K., Grabowski, J.H., Lindquist, N.L., Peterson, C. H., Piehler, M.F., Ridge, J.T., 2017. Oyster reefs as carbon sources and sinks. In: Proceedings of the Royal Society B, 284, p. 1859.

Fuentes, M.M.P.B., Cinner, J.E., 2010. Using expert opinion to prioritize impacts of climate change on sea turtles' nesting grounds. J. Environ. Manag. 91, 2511–2518.

Gordon, T.J., Helmer, O., 1964. Report on a Long-Range Forecasting Study. Rand Corporation Paper P-2982, September.

Grabowski, J.H., Brumbaugh, R.D., Conrad, R.F., Keeler, A.G., Opaluch, J.J., Peterson, C. H., Piehler, M.F., Powers, S.P., Smyth, A.R., 2012. Economic valuation of ecosystem services provided by oyster reefs. Bioscience 62 (10), 900–909.

Hanea, A.M., McBride, M.F., Burgman, M.A., Wintle, B.C., 2018. Classical meets modern in the IDEA protocol for structured expert judgement. J. Risk Res. 21 (4), 417–433.

Hemming, V., Burgman, M.A., Hanea, A.M., McBride, M.F., Wintle, B.C., 2018. A practical guide to structured expert elicitation using the IDEA protocol. Methods in Ecology and Evolution 9 (1), 169–180.

Higgins, C.B., Stephenson, K., Brown, B.L., 2011. Nutrient bioassimilation capacity of aquacultured oysters: quantification of an ecosystem service. J. Environ. Qual. 40 (1), 271–277.

Holsapple, C.W., Joshi, K.D., 2002. Knowledge manipulation activities: results of a Delphi study. Inf. Manag. 39, 477–490.

Humphries, A.T., La Peyre, M.K., 2015. Oyster reef restoration supports increased nekton biomass and potential commercial fishery value. PeerJ 3, e1111. https://doi.org/10.7717/peerj.1111.

Interis, M.G., Petrolia, D.R., 2016. Location, location, habitat: how the value of ecosystem services varies across location and by habitat. Land Econ. 92 (2), 292–307.

Kellogg, M.L., Smyth, A.R., Luckenbach, M.W., Carmichael, R.H., Brown, B.L., Cornwell, J.C., Piehler, M.F., Owens, M.S., Dalrymple, D.J., Higgins, C.B., 2014. Use of oysters to mitigate eutrophication in coastal waters. Estuar. Coast Shelf Sci. 151, 156–168.

Mann, R., Powell, E.N., 2007. Why oyster restoration goals in the Chesapeake Bay are not and probably cannot be achieved. J. Shellfish Res. 26 (4), 905–917.

Martin, T.G., MA, Burgman, Fidler, F., Kuhnert, P.M., Low-Choy, S., McBride, M., Mengersen, K., 2012. Eliciting expert knowledge in conservation science. Conserv. Biol. 26 (1), 29–38.

McBride, M.F., Garnett, S.T., Szabo, J.K., Burbidge, A.H., Butchart, S.H.M., Christidis, L., Dutson, G., H A, Ford, Loyn, R.H., Watson, D.M., Burgman, M.A., 2012. Structured elicitation of expert judgments for threatened species assessment: a case study on a continental scale using email. Methods in Ecology and Evolution 3, 906–920.

McDonald, R.B., RM Moody, KL Heck, Cebrian, J., 2016. Fish, macroinvertebrate and epifaunal communities in shallow coastal lagoons with varying seagrass cover of the northern Gulf of Mexico. Estuar. Coast 39, 718–730.

Meyer, D.L., Townsend, E.C., Thayer, G.W., 1997. Stabilization and erosion control value of oyster cultch for intertidal marsh. Restor. Ecol. 5 (1), 93–99.

Nelson Sella, K.A., Sicius, L., Fuentes, M.M.P.B., 2019. Using expert elicitation to determine the relative impact of coastal modifications on marine turtle nesting grounds. Coast. Manag. 47 (5), 492–506.

O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E., Rakow, T., 2006. Uncertain Judgements: Eliciting Experts' Probabilities. John Wiley & Sons, Ltd, Chichester, UK.

Okoli, C., Pawlowski, S.D., 2004. The Delphi method as a research tool: an example, design considerations and applications. Inf. Manag. 42, 15–29.

Piazza, B.P., Banks, P.D., Peyre, MK La, 2005. The potential for created oyster shell reefs as a sustainable shoreline protection strategy in Louisiana. Restor. Ecol. 13 (3), 499–506.

Piehler, M.F., Smyth, A.R., 2011. Habitat-specific distinctions in estuarine denitrification affect both ecosystem function and services. Ecosphere 2 (1), 1–17.

Scolozzi, R., Morri, E., Santolini, R., 2012. Delphi-based change assessment in ecosystem service values to support strategic spatial planning in Italian landscapes. Ecol. Indicat. 21, 134–144.

Scyphers, S.B., Powers, S.P., Heck Jr., K.L., Byron, D., 2011. Oyster reefs as natural breakwaters mitigate shoreline loss and facilitate fisheries. PloS One 6 (8), e22396. https://doi.org/10.1371/journal.pone.0022396.

Smyth, A.R., Geraldi, N.R., Piehler, M.F., 2013. Oyster-mediated benthic-pelagic coupling modifies nitrogen pools and processes. Mar. Ecol. Prog. Ser. 493, 23–30.

Strand, J., Carson, R.T., Navrud, S., Ortiz-Bobea, A., Vincent, J.R., 2017. Using the Delphi method to value protection of the Amazon rainforest. Ecol. Econ. 131, 475–484.

Sutton-Grier, A.E., Wowk, K., Bamford, H., 2015. Future of our coasts: the potential for natural and hybrid infrastructure to enhance the resilience of our coastal communities, economies and ecosystems. Environ. Sci. Pol. 51, 137–148.

Veen, D., Stoel, D., Zondervan-Zwijnenburg, M., van de Schoot, R., 2017. Proposal for a five-step method to elicit expert judgment. Front. Psychol. 8 (Dec.), 2110.

Weitzman, M.L., 2001. Gamma discounting. Am. Econ. Rev. 99 (1), 260–271.

Wilcox, C., Mallos, N.J., Leonard, G.H., Rodriguez, A., Hardesty, B.D., 2016. Using expert elicitation to estimate the impacts of plastic pollution on marine wildlife. Mar. Pol. 65, 107–114.