Same law, different results: comparative analysis of Endangered Species Act consultations by two federal agencies

Megan Evansen Corresp., [1, 2] , **Ya-Wei Li** Corresp., [2] , **Jacob Malcom** Corresp. [2]

[1]CONS Graduate Program, University of Maryland at College Park, College Park, MD, United States
[2]Endangered Species Conservation, Defenders of Wildlife, Washington, DC, United States

Corresponding Authors: Megan Evansen, Ya-Wei Li, Jacob Malcom
Email address: evansen@umd.edu, YLi@defenders.org, jmalcom@defenders.org

Evaluating how wildlife conservation laws are implemented is critical to determining how best to protect biodiversity. Two agencies, the U.S. Fish and Wildlife Service and National Marine Fisheries Service (FWS and NMFS; Services collectively), are responsible for implementing the U.S. Endangered Species Act (ESA). This creates a "natural experiment" for understanding how implementation of the same law varies between agencies with different histories, cultures, and funding levels. We take advantage of this natural experiment to quantify differences in how FWS and NMFS implement a core component of the ESA, section 7 consultations. The ESA requires federal agencies to consult with the Services if an action an agency proposes might affect ESA-listed species or their habitats. We quantified the quality of consultations by comparing >120 consultations to the requirements laid out in the Services' consultation handbook. These analyses were complemented with in-person interviews of biologists from the Services to help understand how some observed variation arises. We found consultations from NMFS had significantly higher quality scores than those from FWS. A common shortcoming from both agencies, but especially severe for FWS, was the lack of accounting for effects that were previously authorized through consultations. The biologist interviews indicated some discrepancy between how they perceive consultations and the outcomes from our quantitative analysis. Building from these results, we recommend several actions that can improve quality of consultations, such as using a single database to track and integrate previously authorized harm in new analyses, and the careful but more widespread use of programmatic consultations.

Same law, different results: comparative analysis of Endangered Species Act consultations by two federal agencies

Megan Evansen[1,2], Ya-Wei Li[2], and Jacob Malcom[2]

[1]**CONS Graduate Program, 1213 H.J. Patterson Hall, College Park, MD 20740**
[2]**Endangered Species Conservation, Defenders of Wildlife, 1130 17th Street NW, Washington, DC 20036**

Corresponding author: Megan Evansen[1]

Email address: mevansen@defenders.org

ABSTRACT

Evaluating how wildlife conservation laws are implemented is critical to determining how best to protect biodiversity. Two agencies, the U.S. Fish and Wildlife Service and National Marine Fisheries Service (FWS and NMFS; Services collectively), are responsible for implementing the U.S. Endangered Species Act (ESA). This creates a "natural experiment" for understanding how implementation and interpretation of the same law varies between agencies with different histories, cultures, priorities and funding levels. We take advantage of this natural experiment to quantify differences in how FWS and NMFS implement a core component of the ESA, section 7 consultations. The ESA requires federal agencies to consult with the Services if an action an agency proposes might affect ESA-listed species or their habitats. We quantified the quality of consultations by comparing >120 consultations to the requirements laid out in the Services' consultation handbook. These analyses were complemented with in-person interviews of biologists from the Services to help understand how some observed variation arises. Among these consultations, we found those from NMFS had significantly higher quality scores than those from FWS. A common shortcoming from both agencies, but especially severe for FWS, was the lack of accounting for effects that were previously authorized through consultations. The biologist interviews indicated some discrepancy between how they perceive consultations and the outcomes from our quantitative analysis. Building from these results, we recommend several actions that can improve quality of consultations, such as using a single database to track and integrate previously authorized harm in new analyses, and the careful but more widespread use of programmatic consultations.

INTRODUCTION

The U.S. Endangered Species Act (ESA) is considered one of the strongest wildlife laws in the world (Gosnell 2001). Signed into law in 1973 by President Richard Nixon in response to rising concern over the number of species threatened by extinction, the ESA provides over 1,650 U.S. species with protection as of 2017 (USFWS 2017). Today, the ESA remains the primary piece of environmental legislation for protecting imperiled species and recovering them to the point that the law's protections are no longer needed. With such a crucial role, the ESA must be implemented correctly.

Section 7 of the ESA directs federal agencies to use their authorities to conserve listed species, and is a key reason for the law's strength. Under section 7(a)(2), federal agencies are instructed to ensure, in consultation with the U.S. Fish and Wildlife Service (FWS) and the National Marine Fisheries Service (NMFS), that any action authorized, funded, or carried out by such agency ("action agency") is not likely to jeopardize (see Box 1, Glossary) the continued existence of any endangered or threatened species or destroy or adversely modify designated critical habitat. The assessment of these actions by federal agencies and the Services are classified as informal consultations for actions that are deemed not likely to adversely affect listed species or their critical habitat, or formal consultations for those that are likely to adversely affect either. If an action agency concludes not likely to adversely affect, it must request Service concurrence on that finding. If the Service concurs, the consultation is completed. In addition to implementing other programs, e.g., Magnuson-Stevens, the National Wildlife Refuge System, NMFS and FWS share administration of the

ESA and are responsible for consulting with federal agencies on actions affecting listed species under their respective jurisdictions. Generally, NMFS has jurisdiction over marine species while FWS manages terrestrial and freshwater species (USFWS and NOAA 1974), but both Services have jurisdiction over some listed species, such as anadromous salmonids and sea turtles. Action agencies consult with both Services on these joint-jurisdiction species. If done properly, consultations minimize the negative effects of an action and ensure that it does not violate the jeopardy and adverse modification prohibitions.

Box 1: Glossary

Glossary of terms typically used to describe and discuss consultations under section 7(a)(2) of the U.S. Endangered Species Act. The exact legal and policy definitions can be found in the referenced Code of Federal Regulations (CFR) and Handbook sections.

Action All activities or programs of any kind authorized, funded, or carried out, in whole or in part, by Federal agencies in the United States or upon the high seas. [50*CFR§*402.02]

Action agency The federal agency proposing the action.

Biological opinion The document resulting from formal consultation that describes the proposed action, the Service evaluation of the effects of the action, the determination of whether the species' existence is jeopardized or its critical habitat is adversely modified, and any conservation requirements for the action agency. [50*CFR§*402.02, 50*CFR§*402.14(*h*)]

Critical habitat The specific areas and habitats essential to conserving the species. Critical habitat may be designated in areas that are occupied or unoccupied at the time of listing. Occupied habitat must also have "physical or biological features" that require special management considerations or protection. [*ESA§*3(5)(*A*)]

Formal consultation The type of detailed evaluation undertaken for federal actions that are likely to adversely affect one or more ESA-listed species. [50*CFR§*402.02, 50*CFR§*402.14]

Informal consultation The type of detailed evaluation undertaken for federal actions that are not likely to adversely affect one or more ESA-listed species. [50*CFR§*402.02, 50*CFR§*402.13]

Jeopardy (jeopardize) To engage in an action that reasonably would be expected, directly or indirectly, to reduce appreciably the likelihood of both the survival and recovery of a listed species in the wild by reducing the reproduction, numbers, or distribution of that species. [50*CFR§*402.02]

Programmatic consultation A consultation that addresses multiple actions taken by an agency on a program, regional, or other basis. For example, programmatic consultations may cover many different energy development projects within particular Bureau of Land Management lands in a single, landscape-level evaluation. (Handbook, p. xvii)

The Services collaborated to create the Section 7 Handbook to "promote efficiency and nationwide consistency [of consultations] within and between the Services" (USFWS and NMFS 1998). The Handbook guides biologists to ensure consultations are serving their purpose of adequately protecting listed species, for example by specifying required analyses. But the Handbook is a guidance document for a national program and not all details of a consultation are prescribed, allowing enough discretion for variation in consultation quality to arise. Two general observations suggest consultation quality may differ between the Services, which may reduce consultation effectiveness. First, Malcom and Li recently analyzed data on all 88,290 section 7 consultations recorded by FWS between 2008 and early 2015 (Malcom and Li 2015). Among other

results, they found that the duration of the consultations was typically much shorter than the maximum allowed by regulation, with 80% of formal consultations completed within the time limits set by the Handbook. (The proportion of on-time consultations is likely higher because the data do not include information on legitimate "pauses" during consultation; JWM and Y-WL, pers. obs.) In contrast, NMFS consultations are often behind schedule, with only ~30% of formal consultations completed within the required 135-day timeframe (NMFS 2014). One possible explanation for the time difference between the Services that could be problematic for conservation is that FWS may be rushing consultations because the agency has to consult on many more actions but has similar overall funding as NMFS. Second, in reading hundreds of consultation documents, the authors have observed extensive variation in what we loosely refer to as "quality" and "consistency." The variation appears to be structured (e.g., by species or office) rather than random, and our impression is that the largest differences are between the Services. These observations are set against a backdrop of two agencies with different histories, levels of funding, and cultures — often varying by region and office within each Service — that we expect generate the variation (see, e.g., Lowell and Kelly 2016). To our knowledge, there has never been a systematic analysis of these differences in consultation quality.

To evaluate variation in how section 7 is implemented by the Services, we examined the quality of consultations relative to the requirements of the Section 7 Handbook. We expect consultations that follow the requirements of the Handbook are more likely to result in better conservation outcomes—that is, are higher quality—because the Handbook provides the best available description of protections to comply with section 7. We hypothesized that the quality of NMFS consultations was significantly higher than the quality of FWS consultations. To test our hypothesis, we read and scored the quality of > 120 consultations from the Services and conducted interviews with consultation biologists to better understand the basis of variation. We considered completely random sampling, but consultations are often highly context-specific and can vary widely depending on action type, species consulted upon, and other factors. Much like Owen (2012) did in his analysis, we chose a specific subset of consultations to make comparisons between the Services more direct. To control for extraneous sources of variation, we restricted the consultations to those:

1. From Florida, to minimize geographic variation;
2. Focused on sea turtles, to minimize natural history variation of the consulted-on species that could confound analyses;
3. Involving Army Corps of Engineers as the action agency, to maximize the similarity of the types of actions evaluated; and
4. From the period 2008 through mid-2015, to match temporal conditions.

We found significant differences in the quality of both the formal and informal consultations between the Services. The results highlight ways the Services can systematically improve the quality of consultations, in particular, in tracking and analyzing previously authorized take.

METHODS

Consultation Selection

Biological opinions from NMFS consultations are available to the public through their Public Consultation Tracking System (PCTS; https://pcts.nmfs.noaa.gov/pcts-web/homepage.pcts). This database allows users to search for specific consultations or all consultations within specified search parameters. The Tracking And Integrated Logging System (TAILS) is FWS's database for recording consultation data. While PCTS allows users to download consultations in full, TAILS is designed to help coordinate record-keeping between field and regional offices of FWS and does not provide the consultation documents. Instead, the TAILS database offers records of each of the consultations completed by FWS, and interested parties must obtain the consultation documents by other means. TAILS has no public interface, but Malcom and Li (2015) created a web application, the Section 7 Explorer (https://cci-dev.org/shiny/open/section7_explorer/), that allows the public to search for consultations of interest using a number of parameters. The data in the Section 7 Explorer

are updated periodically when FWS provides a new batched data release.

Using PCTS and the Section 7 Explorer, we randomly selected 30 formal and 30 informal consultations from each Service from 2008 to mid-2015. To minimize natural history and geographic variation of the species consulted on by NMFS and FWS, we limited our consultations to those dealing with sea turtles in Florida (green sea turtle [*Chelonia mydas*], loggerhead sea turtle [*Caretta caretta*], Kemp's ridley sea turtle [*Lepidochelys kempii*], leatherback sea turtle [*Dermochelys coriacea*], and hawksbill sea turtle [*Eretmochelys imbricata*]). To minimize confounding variation that could arise if different action agencies were evaluated, we limited consultations to those with the Army Corp of Engineers. We acquired the NMFS consultations directly from PCTS, while those from FWS we acquired through FWS South Florida Field Office's online document library for biological opinions (https://www.fws.gov/verobeach/verobeach_old-dontdelete/sBiologicalOpinion/index.cfm) or through a Freedom of Information Act (FOIA) request. While evaluating the original selection of NMFS formal consultations, we discovered some that did not assess sea turtles in the biological opinion despite search parameters constrained to sea turtles. To account for this discrepancy, we removed those not assessing sea turtles and randomly selected an additional 10 formal NMFS consultations for evaluation from the PCTS database. All of the consultations analyzed in this work are archived at Open Science Framework (OSF) under https://dx.doi.org/10.17605/OSF.IO/KAJUQ.

Evaluation Criteria

We recorded general information for each consultation, such as the start and end dates of the consultation, year it was completed, regional office it was filed through, species of sea turtles concerned, and page length. The full dataset and metadata describing all variables are provided alongside the consultations at OSF (https://dx.doi.org/10.17605/OSF.IO/KAJUQ). Below we describe the scoring methodology, noting that formal and informal consultations required different scoring rubrics because they involve different content. All scoring rubrics are provided in SI Appendix 1 (formal consultations) and SI Appendix 2 (informal consultations). It is important to note that it is not feasible to blind scorers as to the Service that wrote a biological opinion because of the nature of the documents: any familiarity with the consultation process makes the Service immediately apparent. Therefore, the reviewers were not blind to the Service when analyzing quality. To minimize bias, we used a strict set of standards from the section 7 Handbook to analyze quality to the best of our ability. When there was any ambiguity as to the appropriate score, a second reviewer (JWM) would read the consultation in question, then decide on the appropriate score with the primary reviewer (ME).

For formal consultations, we selected the four core sections from the Handbook to score the quality of each biological opinion: "Status of the Species," "Environmental Baseline," "Effects of the Action," and "Cumulative Effects." While not an exhaustive list of biological opinion sections, these four sections contain the bulk of the information and analysis of the species and the proposed action. Each section received a score from 0-5 or 0-2 based on how well they met the specific requirements set out for that section by the Handbook. In developing the scoring system, we found that rating the quality of these core sections of the biological opinion was clear because criteria set by the Handbook allowed for a simple present/absent scoring system. These present/absent scores were summed for each of the four core sections, giving them a maximum possible score of 2 or 5 points. We calculated total quality by summing the scores across all four sections. The overall quality was normalized by calculating the ratio of the summed score to the total points possible for each consultation.

Scoring the informal consultations used a simpler rubric because informal consultations are much shorter, rarely have individual sections, and the Services generally have not prescribed the type of contents that informal consultation documents must contain. We surveyed a selection of informal consultation documents from both Services and several regions, and considered what information Services personnel need to evaluate the effects of actions and to monitor the action after consultation is complete. We identified five criteria to evaluate the quality of informal consultations: mentioning the action, analysis of the action, analysis of the impacted species, mentioning the reason the consultation stayed informal, and including a map of the area affected by the action. These criteria were worth 1 point each, and thus informal consultations received a

quality score from 0-5.

During the preliminary work we noticed the use of "sticker concurrences," in which the FWS South Florida Office record of their analysis consisted only of a sticker of consent applied to the request for concurrence provided to FWS (SI Figure 1). This sticker of approval for the action worked in lieu of a complete informal consultation, and no additional consultation documentation to detail any analysis on the action was supplied.

Statistical Analyses

Our goal was to understand patterns and associations of variation in consultation quality. We used basic summary statistics (e.g., mean, standard deviation) and simple correlations (Pearson's) to describe basic patterns. The analyses proceeded from the broadest scope (factors associated with overall quality, across all consultations) to increasingly detailed analyses of the quality components. We used two basic modeling approaches across this hierarchy: a binomial generalized linear model (GLM; McCullagh and Nelder 1989) on the proportions of total possible points, and ordinal logistic regression (OLR; Kleinbaum and Klein 2010) of the individual quality component scores. We considered seven variables that we thought were most likely to affect consultation quality: the Service that performed the consultation, whether the consultation was formal or informal, the year the consultation took place, the species of turtle assessed, the type of action assessed, and whether the consultation was part of a programmatic consultation. We incorporated these variables into a set of nine candidate models for the analysis of overall quality using the GLM (Table 1, "GLM binom."). Our global model (Model 1) contained all seven variables. We also considered that the particular office within the Service might be an important predictor of consultation quality. However, given that our focus is on the potential differences between the Services and that the offices are nested within the Services, the office variable was not included in our candidate model set. Because of the fundamental differences between formal and informal consultations and the difference in total possible score, we calculated the response variable as the proportion of possible points for each consultation. When we investigate within formal and informal consultations, we used reduced candidate model sets, dropping variables that were not useful. This meant removing the formal consultation variable from formal analyses, and the formal and programmatic variables from the informal analyses.

To evaluate the quality components, we used a set of three candidate ordinal regression models (Table 1, "Ord. regress") with random effects for the consultation document in which the components were nested. While programmatic consultation was an important predictor of quality in the overall analysis, the Hessian was singular (presumably because of the lack of NMFS programmatic consultations) for the components and we were not able to include programmatic as a variable in these analyses. In lieu of more complex analysis, we evaluated simple summary statistics to investigate the role of programmatic consultations in shifting quality scores. We used package `ordinal` (Christensen 2015) for the ordinal regression.

We used Akaike's Information Criterion adjusted for small sample sizes (AIC_C) for model selection (Anderson and Burnham 2002) using the AICcmodavg package (Mazerolle 2011). All analyses were done in R 3.3 (R Core Team 2016) and are available as a package vignette in the project's OSF repository (https://dx.doi.org/10.17605/OSF.IO/KAJUQ).

Biologist Interviews

To better understand the consultation process, one of the authors (ME) conducted semi- structured interviews (see, e.g., Pienaar 2015 for an example of this) with biologists from both Services and one biologist from the Florida Fish and Wildlife Conservation Commission who works closely with the Services. Interviews were conducted concurrent with our scoring of the consultations, in August 2015, and the interview questions were based on our understanding of the Handbook and preliminary examination of the consultations we reviewed. We asked the same questions of all interviewees regarding their views on the consultation process and how well consultations serve their intended purpose (SI Appendix 3). We interviewed all biologists under the condition of anonymity. Although the sample size is too small for statistical analysis, we reviewed and scored the notes from the interviews to summarize recurring themes.

RESULTS

We retrieved, read, and scored 123 consultations from the two Services (Table 2). Summary statistics for both formal and informal consultations are provided in Table 3. On average, the analyzed consultations assessed the effects of the action on seven species. Formal consultations ranged in length from 1 page to 120 pages and took over a year on average to complete. Of the core quality sections evaluated, 'Status of the Species' was by far the longest, with an average of 18.65 pages. We noted that this section often contained extensive extraneous material that was not relevant to the species' life history in the area of the action, nor was the information relevant to the effects of the action. In our random sample of FWS informal consultations, only one had the sticker concurrence that we observed in the preliminary work.

Overall Consultation Quality

Model 9 was the best supported among our candidate model set for the quality sections of consultations (Table 4). This model, which included all predictors except action type, indicated that a consultation done by NMFS was 1.40 times (95% CI = 1.25 - 1.57; Figure 1a) as likely to receive a positive score for quality components as a consultation done by FWS; FWS's programmatic consultations provided a significant quality boost (OR = 1.35; 95% CI = 1.17 - 1.56); but formal consultations were about as likely (OR = 1.0; 95% CI = 0.89 – 1.13; Figure 1b) to score positively as informal consultations (Table 5). We found that the duration of consultations was positively associated with overall quality in a simple univariate analysis ($r = 0.20$; $p = 1.04e^{-6}$), but disappeared in the multivariate analysis. Similarly, the length (in pages) of consultations was also correlated with quality in a univariate analysis (r = 0.2, p = 0.0037). However, after accounting for the Service performing the consultation and for programmatic consultations in a binomial GLM, there was no relationship ($z = 1.024$, $p = 0.306$).

Quality Components

We next examined the sources of variation in the components of overall consultation quality. The only component of formal consultations that exhibited a strong signal with any predictors was the Environmental Baseline, for which Service was a significant and strong predictor of quality ($z = 5.3993$, $p = 6.691e\text{-}08$; $OR_{NMFS} = 2.6e^4$ [95% CI = $6.5e^2 – 1.1e^6$]). These patterns are readily visible (Figure 2), and suggest that NMFS consultations tended to score better even though there were few statistically supported differences. For the Environmental Baseline section, NMFS consultations tended to include previous consultations in the action area and discuss critical habitat or lack thereof, neither of which were consistently present in FWS consultations.

Most of the quality components of informal consultations were relatively homogenous (Figure 3), with two exceptions. The analysis of the action and the reason the consultation was informal were significantly (at a nominal $\alpha = 0.05$) associated with the duration of consultation: the longer the informal consultation, generally, the more likely these components were included. Second, although not required by the Consultation Handbook, half of NMFS informal consultations included a map of the proposed action but only 15% of FWS informal consultations did.

Interviews

We interviewed seven biologists from FWS and NMFS who consult on section 7 actions and tallied their responses to our questions (Table 6; full response notes in SI Appendix 4). When asked how the consultation process could be improved, most biologists (6/7) mentioned they found the process frustrating and many stated that they were overwhelmed with work. One biologist pointed to the fear of possible litigation resulting from shorter consultations as a reason for the overly comprehensive and highly time-consuming consultations that are currently the norm. Five of seven biologists also favored expanding the use of consultation "keys," which are designed to help the biologists improve the timing and consistency of consultations (see, e.g., http://www.fws.gov/panamacity/resources/WoodStorkConsultationKey.pdf; SI Appendix 5) when

appropriate for a species or on a case-by-case basis. All biologists interviewed except one mentioned that they keep a record of cumulative incidental take to the best of their ability. The method of recording authorized take varied from notes kept on a whiteboard to Excel spreadsheets. However, only three consultations (all from NMFS) received a positive score for incorporating previously authorized take in the analysis of the effects of the current action on sea turtle populations.

DISCUSSION

The ESA is considered the strongest national wildlife protection law in the world, and section 7 is a key reason for this strength. The quality of section 7 consultations can alter conservation outcomes because the protections afforded by the section can only be realized if the scientific and regulatory analyses are robust. Despite the importance of consistently high-quality consultations, no analyses have critically evaluated the strengths and weaknesses of these regulatory documents. Our analysis is a first step for understanding the quality of past consultations to inform and improve future consultations. Across all 123 consultations evaluated, we found that quality varied significantly between the Services and our hypothesis that the quality of NMFS consultations is higher than FWS consultations was supported. In combination with the biologist interviews, which shed light on some of the causes of variation, our results suggest ways that consultations can be improved.

Quality Differences

The quality scores of NMFS consultations were significantly higher than those of FWS for both formal and informal consultations, consistent with our hypothesis. This is also consistent with the findings of Lowell and Kelly (2016), who found NMFS scored higher than FWS in three of seven metrics characterizing the use of "Best Available Science" in recovery plans, lawsuits, listing decisions, and literature cited in biological opinions. The ultimate cause of the difference is unclear, but one likely explanation comes from our interviews. FWS biologists in particular spoke repeatedly about the lack of time and resources for an ever-increasing consultation workload. This sentiment reflects the broad-scale funding shortfall that the FWS endangered species program faces: it receives about the same amount of funding as the Office of Protected Resources at NMFS, even though Ecological Services within FWS is responsible for 15 times as many ESA-listed species (Lowell and Kelly 2016). We do not have data on how the Services allocate funding to consultations versus other endangered species program components, such as listing and recovery, but spending per consultation is likely much lower for FWS.

Our scoring of the individual sections of biological opinions allows us to better understand why FWS consultations are lower quality and where both Services deviate from the expectations of the Handbook. The Environmental Baseline section of consultations we evaluated consistently earned a score less than the maximum possible (= 5 points) because previously authorized incidental take in the action area was rarely analyzed. The lack of this analysis was problematic for both Services, but FWS scored significantly lower ($\bar{x}_{FWS}$ = 2.44) in the Environmental Baseline than NMFS ($\bar{x}_{NMFS}$ = 3.59) because the take analysis was missing from all prior consultations in the action area we evaluated. This may seem a minor point, but the lack of this analysis is one of the more pernicious problems of implementing the ESA (Owen 2012). The occurrence of hundreds or thousands of small actions can too easily result in "death by a thousand cuts," whereby individual actions are insignificant for the species, but the cumulative effects across many actions may severely damage their populations (USFWS 2012). A 2009 Government Accountability Office report on FWS's implementation of the ESA highlighted this concern and recommended that the Services track authorized take across a species' entire range to better inform consultations (GAO 2009). The three consultations that included an analysis of previously authorized take were all done by NMFS, enhancing the quality difference between the Services for this core section. However, it is worth noting that FWS's programmatic consultation for beach work across Florida (Activity Code 41910-2010-F-284) did list previous formal consultations. Unfortunately, those data were not analyzed in the consultation and there was no evidence they played a role in the Environmental Baseline or the Effects Analysis. Why previously authorized take in the action area is not

analyzed is unclear, especially in light of the interviews in which many biologists stated that they personally track cumulative take. Future work should investigate the disconnect between the information that Services biologists record and the information used in consultations.

The Handbook requires certain components for each section. Unfortunately, several sections of many FWS consultations consisted only of the boilerplate language from the Handbook and little or no analysis, which lowered FWS scores. This was particularly true of the Cumulative Effects section of FWS consultations, which often mention the obligation to "include the effects of future State, tribal, local or private actions that are reasonably certain to occur," then simply stated that there would be no cumulative effects. In contrast, most NMFS consultations more thoroughly analyzed the cumulative effects, which are critical to understanding the effects on species recovery.

The Handbook guidance for informal consultations is much less prescriptive than for formal consultations, but our analysis shows FWS lagging behind NMFS for this large set of consultations. Three components — the analysis of the action, the species analysis, and a map of the action area — were systematically missing or insufficient in the informal FWS consultations we reviewed. On one hand, we recognize that detailed analysis of actions covered by an informal consultation is unlikely to benefit ESA-listed species because the main purpose of those consultations is to determine if a more detailed formal consultation is needed. But the trade-off is that some of the most important components of the administrative record are missing. Perhaps the most obvious example of this missing component comes from the use of "sticker" concurrences, observed both in our preliminary work and in one randomly sampled informal consultation. While these stickers may save time, they provide no record of why FWS approved the action, which is critical to understanding whether FWS is properly implementing the ESA. In contrast, all informal consultations from NMFS explained why the consultation was informal. The shortcomings of FWS informal consultations can likely be explained by the resource constraints discussed above, but highlight the need for the agency to critically evaluate whether it has sacrificed some conservation in the name of efficiency.

Consultation Efficiency

High quality consultations are essential to properly implementing the ESA, but there is also a need for efficiency. Ideally, the Services should commit to spending enough time on each consultation to maximize the conservation benefit to a listed species across its entire range. Any additional negotiation with project proponents is inefficient, taking resources away from other tasks that could deliver greater conservation benefits. Converse and colleagues (2011) used a decision-analytic approach to identify a point of diminishing returns for bull trout (*Salvelinus confluentus*) consultations in an FWS field office with a global optimum in mind. Such an analysis of the optimal allocation of effort for FWS and NMFS consultations evaluated here is beyond the scope of the present work. Instead, we focus on efficiencies— and potential pitfalls of efficient approaches — indicated by our results.

Programmatic consultations are one promising way to improve consultation efficiency. The effects analysis should provide a better description of cumulative effects because many planned or potential projects within a program are evaluated together rather than individually. We expect that when the cumulative impacts are properly acknowledged, the assessment of jeopardy or adverse modification is more likely to reflect real-world conditions. Another benefit is that because the overall program has already been evaluated, the consultations for future individual projects are faster and can contain less analysis. Malcom and Li (2015) found that project-level consultations that tiered off of a program-level consultation were completed nearly three times faster than the average standard consultation. In the set of consultations we evaluated, the single FWS program-level programmatic consultation for beach renourishment across Florida was a "tide that raised all boats." Whether this is an outlier or representative of programmatic consultations in general is unclear, but deserves further investigation. The project-level programmatic consultations that tiered off of the program-level programmatic consultation "inherited" the (generally) high scores of the program-level consultation and significantly increased the quality of FWS consultations. But the converse is also possible: low-quality program-level programmatic consultations would mean that tiered consultations inherit low-quality analyses that would likely lead to poor conservation outcomes. While the results from this set of consultations are

promising, the Services need to continually evaluate their programmatic consultations to ensure that the speed benefits of these consultations do not overshadow the need for high-quality analyses.

Our interviews with biologists from the Services may be preliminary, but provided important context for interpreting the results and indicated other possibilities for improving consultation efficiency. The lack of consistency among offices and between Services was frequently mentioned as a frustrating aspect of the consultation process during the interviews. The differing approaches to consultations can be difficult for action agencies as well, who can see the approval of a project depend largely on the consulting office (Y-WL and JWM, pers. obs.). One possible solution that we did not test is the use of consultation keys, as have been developed for Army Corps of Engineers consultations for a few species, including wood storks (*Mycteria americana*) and indigo snakes (*Drymarchon couperi*). The Services use these documents to promote appropriate standards for certain construction activities. Creating similar documents for other frequently-consulted species may streamline consultations and increase inter-office and inter-Service consistency. The use of consultation keys would also increase the transparency of the consultation process, making it easier for action agencies or their applicants to plan their projects.

Last, we note one particular aspect of consultations that was not amenable to quantitative analysis but suggests efficiency improvements: inclusion of extensive material seemingly irrelevant to evaluating the effects of the action. For example, several consultations we reviewed included >20 pages of information on red knots (*Calidris canutus*), of which one paragraph was relevant to evaluating the action (JWM, pers. obs.). Including such inconsequential background information requires additional time not only for Services' biologists, but also for the action agency or their applicants who read the opinion. By way of explanation, one FWS biologist mentioned that such information was included to buffer against any potential legal action, ensuring all "bases are covered." However, this approach conflates "more" with "better" — the added time and cost does not always produce commensurate benefits for legal defensibility or conservation (Restani and Marzluff 2002). We encourage the Services to critically evaluate the information in biological opinions, and exclude irrelevant material. The Recovery Enhancement Vision (REV) project being developed by FWS at this time (SI Appendix 6) can help with this extraneous information problem. One component of REV is a single, continually updated Species Status Assessment (SSA) for each ESA-listed species, which would be incorporated by reference in consultations, conservation permits, five-year reviews, and other aspects of ESA implementation (SI Appendix 7). Widespread adoption of SSAs would improve efficiency and, because they should include an analysis of previously authorized take, improve the effectiveness of section 7 consultations.

Policy Recommendations

Our analyses shed new light on how the Services implement section 7 consultations. Do the consultations that we selected reflect all consultations, nationwide? Perhaps not, but the results are sufficient to make two main policy recommendations for the Services to implement:

1. *Develop and require the use of a single database for recording and querying authorized take.* A centralized take database was recommended by the GAO seven years ago (GAO 2009) but has not yet been implemented by the Services. The component most commonly missing from consultations we reviewed was an analysis of previously authorized take in the action area. This is not surprising because the Services lack a unified way for their biologists to record authorized take, much less to tally previously authorized take to use in the jeopardy and adverse modification analyses. Implementing this recommendation would dramatically improve the quality of the Environmental Baseline analysis of consultations. In turn, we expect better conservation outcomes for consulted-on species. Beyond consultations, an authorized take database would be invaluable for informing ESA-required five-year status reviews, such that harmful effects from consultations can be compared to beneficial effects from conservation activities.
2. *Establish a systematic review protocol to ensure that programmatic consultations, which can increase efficiency, do not reduce the effectiveness of consultation.* Programmatic program-level consultations can increase consultation effectiveness and efficiency – in theory – but the Services must ensure that the quality of project-level consultations is not sacrificed. In our results, the programmatic consultation was the "rising tide that lifted all boats." Ensuring that other and future programmatic consultations are similarly well-crafted can result in high quality, consistently- implemented consultations. The Services have expressed an interest in increasing the use

of programmatic consultations, but such an increase must formally guard against a loss of effectiveness. Regular reviews at the field office, regional, and national levels, guided by a robust "checklist" of effectiveness measures, should be instated as part of an expansion of using programmatic consultations.

In addition to the differences we found in our analyses, we observed more variation in consultations than we expected. If we had chosen a wider selection of consultations then this variation would have only increased. This highlights the need to promote standardization as a means of improving the efficiency and effectiveness of consultations. One simple and transparent way to improve consistency that we did not test is for the Services to develop and use consultation keys. Not every species and every type of action is amenable to consultation keys, but their use could significantly improve the parts of consultations where keys are possible. To reduce the rote workload for consultation biologists and consulting agencies, the Services should transition to referencing SSAs in consultations, which dovetails with FWS's current REV and SSA programs. Improving efficiency through standardization should not mean cutting corners, however. The informal concurrence stickers are a form of standardization but, as currently used, do not provide an adequate record of why decisions were made. They may be sufficient if modified slightly, such as by adding simple check boxes and short note fields to indicate the reason a consultation qualified as informal.

We expect that implementing these recommendations would significantly improve the conservation benefit conferred by section 7 consultations and clarity for those engaged in the process. We also think that these recommendations can help reduce the workload for biologists. By improving the quality of the consultations through these methods, the Services can work toward improving the effectiveness of the ESA as a whole.

ACKNOWLEDGMENTS

We thank the personnel from the Florida offices of the U.S. Fish and Wildlife Service; the St. Petersburg office of the National Marine Fisheries Service; and the Florida Fish and Wildlife Conservation Commission for their work on consultations and for the insights they provided us during this project. We also thank J. Clarke and R. Dreher for helpful feedback on earlier drafts.

LITERATURE CITED

Burnham KP, Anderson DR. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer.

Christensen, R. H. B. 2015. ordinal - Regression Models for Ordinal Data. R package version 2015.6-28. http://www.cran.r-project.org/package=ordinal/.

Converse SJ, Shelley KJ, Morey S, Chan J, LaTier A, Scafidi C, Crouse DT, Runge MC. 2011. A Decision-Analytic Approach to the Optimal Allocation of Resources for Endangered Species Consultation. Biological Conservation 144 (1): 319–29. doi:10.1016/j.biocon.2010.09.009.

Government Accounting Office (GAO). 2009. Endangered Species Act: The U.S. Fish and Wildlife Service Has Incomplete Information about Effects on Listed Species from Section 7 Consultations. http://www.gao.gov/assets/290/289991.pdf.

Gosnell, H. 2001. Section 7 of the Endangered Species Act and the Art of Compromise: The Evolution of a Reasonable and Prudent Alternative for the Animas-La Plata Project. Natural Resources Journal 41: 561.

Kleinbaum DG, Klein M. 2010. Logistic regression: A self-learning text. New York: Springer.

Lowell N, Kelly RP. 2016. Evaluating Agency Use of 'best Available Science' under the United States Endangered Species Act. Biological Conservation 196 (April): 53–59. doi:10.1016/j.biocon.2016.02.003.

Malcom JW, Li Y-W. 2015. Data Contradict Common Perceptions about a Controversial Provision of the US Endangered Species Act. Proceedings of the National Academy of Sciences 112 (52): 15844–49. doi:10.1073/pnas.1516938112.

Mazerolle MJ. 2011. AICcmodavg: model selection and multimodel inference based on (Q) AIC (c). R package version 1.

McCullagh P, Nelder JA. 1989. Generalized linear models. Second edition. Chapman and Hall/CRC, Boca Raton, Florida, USA. 532pp.

National Marine Fisheries Service (NMFS). 2014. FY 2014 Budget Summary. National Marine Fisheries Service, Washington, D.C. http://www.corporateservices.noaa.gov/nbo/fy14_bluebook/FINALnoaaBlueBook_2014_Web_Full.pdf.

Pienaar, E.F., Kreye, M.M. & Jacobs, C. Conflicts between Cattlemen and the Florida Panther: Insights and Policy Recommendations from Interviews with Florida Cattlemen. Hum Ecol (2015) 43: 577. doi:10.1007/s10745-015-9765-x.

Owen, D. 2012. Critical Habitat and the Challenge of Regulating Small Harms. Florida Law Review 64: 141.

R Core Team. 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

Restani M, Marzluff JM. 2002. Funding Extinction? Biological Needs and Political Realities in the Allocation of Resources to Endangered Species Recovery. BioScience 52(2): 169–77. doi:10.1641/0006-3568(2002)052[0169:FEBNAP]2.0.CO;2.

U.S. Fish and Wildlife Service (USFWS). 2012. Endangered and Threatened Wildlife and Plants; Designation of Revised Critical Habitat for the Northern Spotted Owl. Federal Register 77 (46). https://www.fws.gov/oregonfwo/species/Data/NorthernSpottedOwl/Documents/NSO-FinalCH_Rule21Nov2012.pdf.

———. 2017. Listed Species Summary (Boxscore). Environmental Conservation Online System. Accessed 2017-06-24. http://ecos.fws.gov/ecp0/reports/box-score-report.

USFWS, National Marine Fisheries Service (NMFS). 1998. Endangered Species Consultation Handbook. U.S. Fish & Wildlife Service and National Marine Fisheries Service. http://www.blm.gov/or/esa/files/ESA_Section7_Consultation_Handbook_1998.pdf.

———. 1974. Memorandum of Understanding between the U.S. Fish and Wildlife Service (DOI) and the National Marine Fisheries Service (NOAA/DOC) Regarding Jurisdictional Responsibilities and Listing Procedures under the Endangered Species Act of 1973. http://www.nmfs.noaa.gov/pr/laws/esa/mou usfws.pdf.

TABLES

Table 1. Candidate models evaluated for predicting overall consultation quality and conservation action specificity.

Model Type	Model Num.	Predictors
GLM Binom*	1	Service + Formal + Year + Action_type + Programmatic + total_duration
	2	Service + Formal + Year + Programmatic + total_duration
	3	Service + Formal + Year + Action_type + total_duration
	4	Service + Formal + Year + total_duration
	5	Service + Formal
	6	Service
	7	Formal
	8	total_duration
	9	Service + Formal + Programmatic + total_duration
Ord. regress.**	1	Service + Year + (1|consultation_ID)
	2	Service + (1|consultation_ID)
	3	Year + (1|consultation_ID)
	4	Programmatic

* Binomial logistic generalized linear model
** Ordinal logistical regression
*** The notation "(1|var)" indicates a random effects variable

Table 2. Number of consultations evaluated for each Service, by consultation type.

	FWS	NMFS
Informal	25	30
Formal	30	38
Total	55	68

Table 3. Summary statistics across all consultations

Group	Variable	Mean	Min	Max	SD	N*
Formal consultations	Length (pages)	34.6	1	120	21.1	284
	Duration (days)	371.5	6	1691	320.2	340
	No. of species (total)	7	4	18	3.6	324
	No. of References	164.3	1	434	121.4	330
	Species Status length (pages)	18.7	0	67	12.5	325
	Baseline length (pages)	6.7	0	23	4.7	318
	Effects length (pages)	5.4	0	15.5	3.9	303
	Cumulative Effects length (pages)	0.7	0	1.5	0.3	298
	CR**	0.9	0	1	0.3	292
	CM**	0.5	0	1	0.5	272
	RPM**	0.8	0	1	0.4	287
Informal Consultations	Duration (days)	163	0	1227	223.3	260
	No. of species	7.0	1	49	6.0	265
	Construction Conditions	0.7	0	1	0.4	264

* Numbers are based on individual turtle species per consultation because the jeopardy and adverse modification conclusion is made on per-species basis for an action. ** CR = Conservation Recommendations made by the Services; CM = Conservation Measures proposed by the action agency; RPM = Reasonable and Prudent Measures to minimize the amount of take resulting from an action

Table 4. Model selection results for overall quality across all FWS and NMFS consultations

Model	K	AICc	Delta AICc	Model Likelihood	Akaike Weight	Log Likelihood	Cum. Wt.
Mod9	5	1544.5	0.00	1.00	0.71	-767.18	0.71
Mod2	6	1546.3	1.79	0.41	0.29	-767.05	1.00
Mod1	14	1558.8	14.33	0.00	0.00	-765.03	1.00
Mod4	5	1561.4	16.90	0.00	0.00	-775.63	1.00
Mod3	13	1571.0	26.51	0.00	0.00	-772.17	1.00
Mod8	2	1574.5	30.08	0.00	0.00	-785.26	1.00
Mod5	4	1601.7	57.28	0.00	0.00	-796.84	1.00
Mod6	2	1607.4	62.94	0.00	0.00	-801.69	1.00
Mod7	2	1628.1	83.65	0.00	0.00	-812.05	1.00

Table 5. Odds ratios (OR), confidence intervals, and parameter statistics for model 9, the best-supported candidate set for predicting overall consultation quality.

	OR	LCL (2.5%)	UCL (97.5%)	Model z-value	p-value
(Intercept)	5.54E-01	4.93E-01	6.23E-01	-9.883	4.94E-23
Service (NMFS)	1.40	1.25	1.57	5.689	1.28E-08
Formal (yes)	1.00	0.89	1.13	0.042	9.66E-01
Programmatic (yes)	1.36	1.18	1.57	4.202	2.64E-05
total duration	1.00	1.00	1.00	1.454	1.46E-01

Table 6. Responses to a selected sample of interview questions asked of FWS/NMFS biologists.

Biologist	Favor consultation keys	Often encounter scientific uncertainty	Tally cumulative take	Frequently reference section 7 Handbook	Favor publicly available consultations	Suggestions for improvement
1	In some cases	No	Yes	Yes	Yes	Inter-office consistency
2	Yes	No	Yes	No	Yes	None
3	No	No	Yes	Variable	Yes	Inter-office consistency
4	Yes	Rarely, assume species is present	Yes	No	Yes	Intra- and inter-office consistency
5	In some cases	Rarely, assume species is present	Makes an attempt	Yes	Yes	BiOp streamlining
6	In some cases	No	Yes	Yes	Yes	Inter-office consistency
7	No, too nuanced	Yes, defer to species	No - too difficult	No	Yes	Improve efficiency

FIGURES

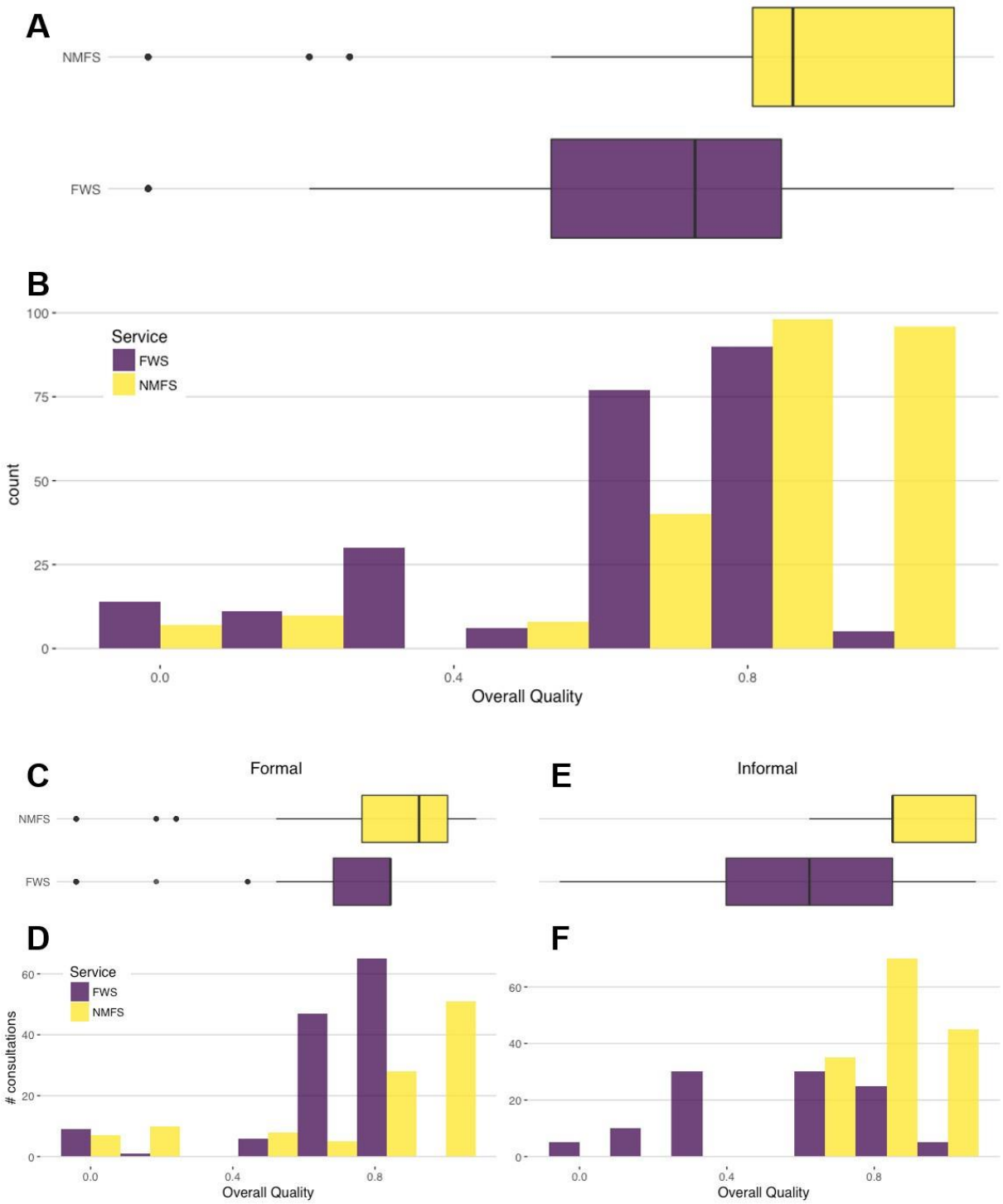


Figure 1. The quality scores for NMFS consultations were higher on average than the scores for FWS consultations across all consultations (A), formal consultations (B), and informal consultations (C). The overall score for each consultation is the sum of points scored divided by the sum of points possible. *Top panel:* Histogram and boxplots of all consultations (formal and informal, including programmatic consultations) for each Service. *Bottom panel:* Overall scores plotted by Service for formal and informal consultations separately.

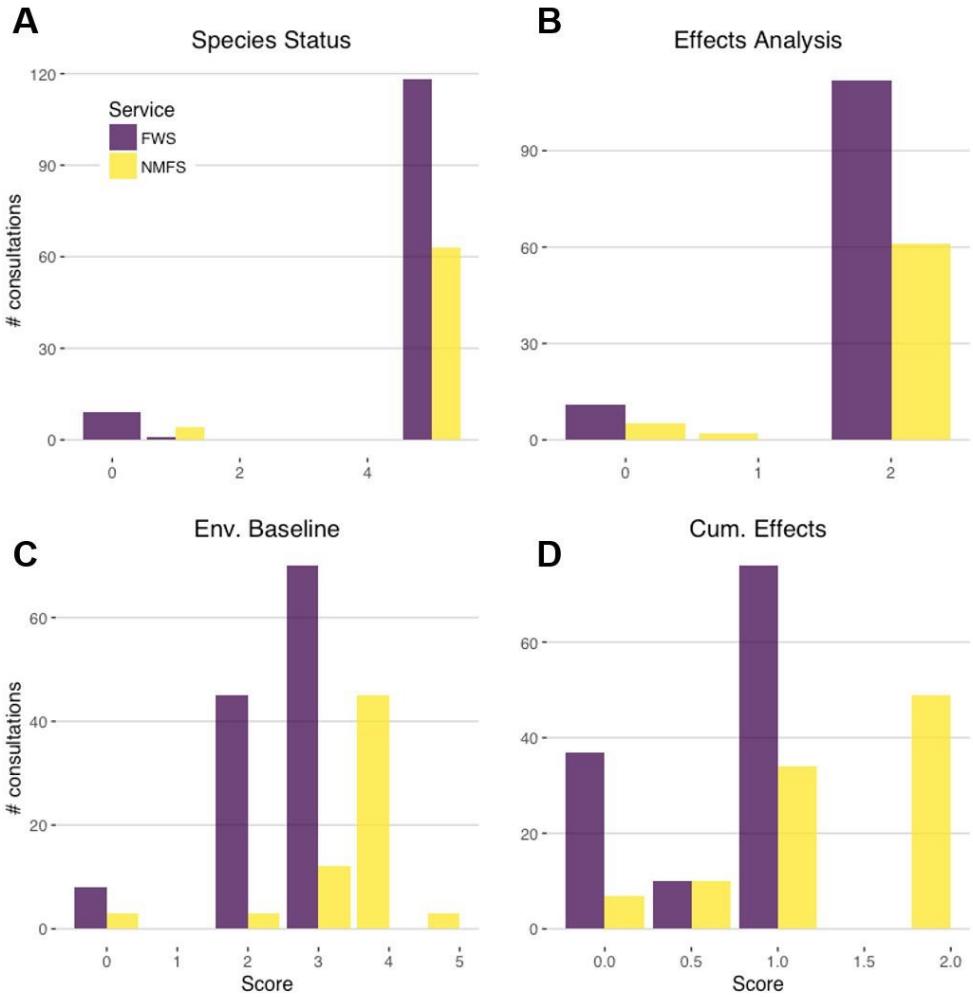


Figure 2. Individual components of consultations from NMFS tended to have higher scores than those from FWS. However, only component with a strong statistical signal for differences between the Services was the Environmental Baseline (z = 5.3993, p = 6.691e-08; OR_{NMFS} = 2.6e^4 [95% CI = 6.5e^2 – 1.1e^6]). The scores are the raw scores for formal consultation components, rather than a proportion as in the overall scores in Figure 1.

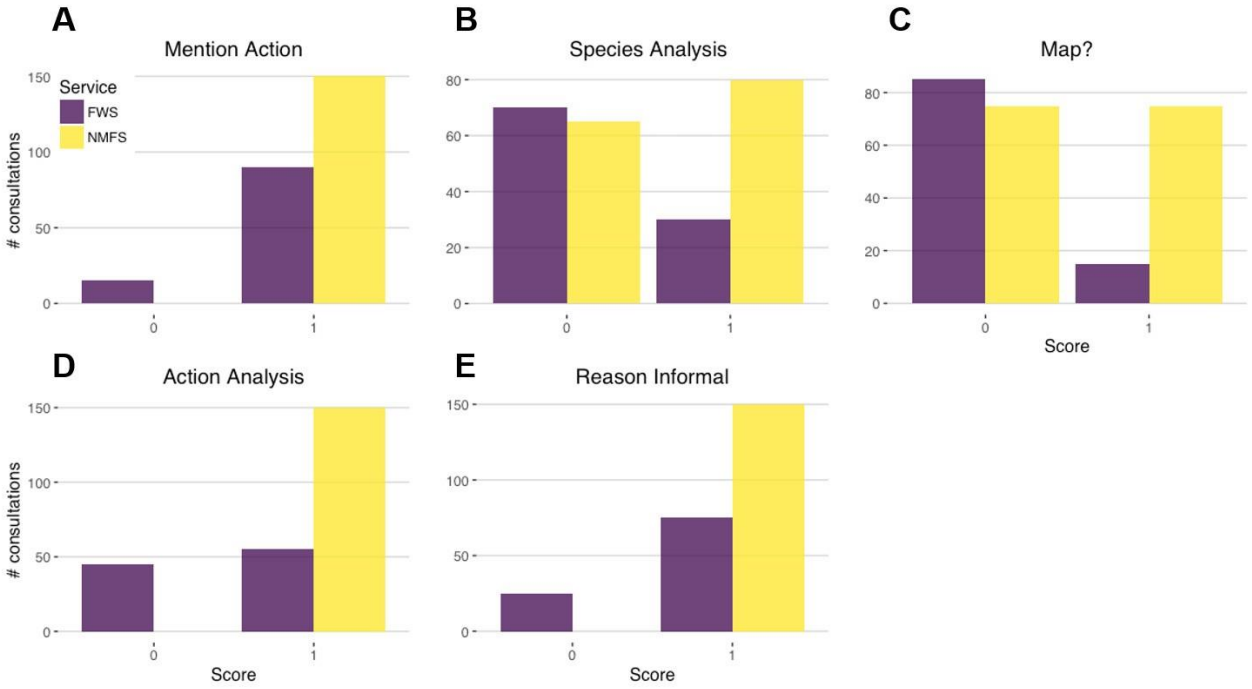
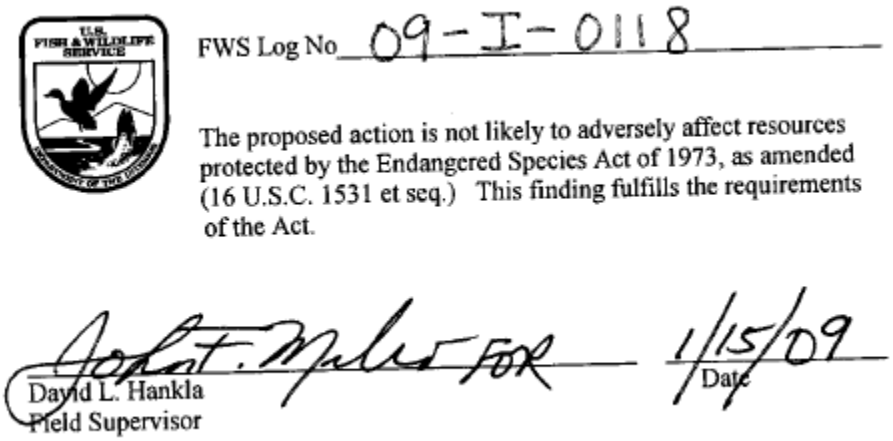


Figure 3. Informal consultations from NMFS tended to have more information and therefore tended to have higher scores than those from FWS, but few differences were statistically significant. The components of informal consultation quality scores were binary (presence/absence) in the consultations.

SI FIGURE 1: INFORMAL STICKER CONCURRENCE



U.S. FISH & WILDLIFE SERVICE

FWS Log No 09-I-0118

The proposed action is not likely to adversely affect resources protected by the Endangered Species Act of 1973, as amended (16 U.S.C. 1531 et seq.) This finding fulfills the requirements of the Act.

FOR 1/15/09

David L. Hankla
Field Supervisor

Date

Complete informal consultation included in Open Science Framework archive at https://dx.doi.org/10.17605/OSF.IO/KAJUQ. Note that there is no accompanying analysis to clarify why this informal consultation was found not likely to adversely affect the species or any listed critical habitat.

SI APPENDIX 1: SCORING RUBRIC FOR FORMAL ESA SECTION 7 CONSULTATIONS

Environmental Baseline (EB) Quality (Total Points: 5)

1. Does the EB address the status of the species in the action area? (1)
2. Is there a mention of past/ongoing threats to the species in the action area? (1)
3. Does the EB take past consultations in the action area into consideration? (1)
4. Is there mention of critical habitat (or lack thereof) for the species? Does said critical habitat overlap with the action area? (1)
5. Does the baseline include State, tribal, local and private actions already affecting the species that will occur contemporaneously with the consultation in progress, as per the handbook? (1)

Effects of the Action Quality (Total Points: 2)

1. There is a clear and defined cause and effect analysis of the action. (1)
2. The consultation gives an explanation as to if and how said action will negatively affect sea turtles. (1)

Species Status Quality (Total Points: 5)

1. Does the consultation adequately describe the species and its habitat/critical habitat? (1)
2. Is the life history of the species addressed? (1)
3. Is there a detailed demographic analysis (if available for the species), including population size, variability and stability? (1)
4. Is the status and distribution of the species addressed, including reasons for listing? (1)
5. Is there an analysis of the species/critical habitat likely to be affected by the action? (1)

Cumulative Effects Quality (Total Points: 2)

1. Does the consultation consider the likelihood of the species to be able to recover? (1)
2. Does the consultation consider the effects of *future* State, tribal, local or private actions that are reasonably certain to occur, as per the handbook? (1)

SI APPENDIX 2: SCORING RUBRIC FOR INFORMAL ESA SECTION 7 CONSULTATIONS

Informal Criteria Baseline (Total Points: 5)

1. Mentions the action (1)
2. Some analysis of the action (1)
3. Some analysis of the impacted species (1)
4. Reason the consultation stayed informal is mentioned (1)
5. Map of the area affected by the action (1)

SI APPENDIX 3: INTERVIEW QUESTIONS FOR FISHA ND WILDLIFE SERVICE, NATIONAL MARINE FISHERIES SERVICE, AND FLORIDA FISH AND WILDLIFE CONSERVATION COMMISSION BIOLOGISTS

1. Can you tell me a bit about how the consultation process usually begins for you?
2. How frequently do you work on consultation? Has this number increased or decreased in recent years? Why might that be so?
3. How common is it to ask the action agency to provide more information on the action?
4. Have you seen a change over time in the way consultations are completed?
5. The number of consultations for FWS in Florida has been steadily decreasing since 2008 (according to the TAILS database there were 1099 in 2008 vs. 347 in 2014). Do you have an impression of how often you aren't consulted on things?
6. Is there a consultation key for sea turtles, similar to the FWS Wood Stork Consultation Key? If not, is this something the Service would consider doing? Would this be an improvement to the process? Would you be in favor of a more standardized way to approach the consultation process? (Keys, a standardized ITP, etc.)
7. Can you explain the process of going through the literature and files on hand to satisfy the "best possible science" condition?
8. How do you exercise precaution when dealing with scientific uncertainty surrounding the effects of an action on a species/critical habitat? How much benefit of the doubt do you give to the species? Does it differ depending on the situation? Is this an issue you deal with on a regular basis?
9. How much time do you spend on the average consultation? FWS TAILS database says the average days for approval for formal consultations is 89 (13 for informal) days. Does that seem right?
10. Is pervious take ever tallied (formally or informally) to get a sense of how much has been done to a species over time? In your view, would this be a feasible/helpful thing to implement?
11. How often do you consult the section 7 Handbook?
12. Do you ever get requests for re-initiation of consultations?
13. NMFS is taking the lead on the revision of the handbook this year. What would you like to see in the revision? In your opinion, is there something that should be clarified?
14. What is your opinion on making all of the final documents publicly available (NMFS has PCTS, Vero Beach has the formal consultations online but not the informal documents)?
15. Where is there the most room for improvement in the consultation process? Does it work well as is?

SI APPENDIX 4: INTERVIEW RESPONSES

Included in Open Science Framework archive at https://dx.doi.org/10.17605/OSF.IO/KAJUQ

SI APPENDIX 5: WOOD STORK CONSULTATION KEY

Included in Open Science Framework archive at https://dx.doi.org/10.17605/OSF.IO/KAJUQ

SI APPENDIX 6: RECOVERY ENHANCEMENT VISION PRESENTATION

Included in Open Science Framework archive at https://dx.doi.org/10.17605/OSF.IO/KAJUQ

SI APPENDIX 7: SPECIES STATUS ASSESSMENT PRESENTATION

Included in Open Science Framework archive at https://dx.doi.org/10.17605/OSF.IO/KAJUQ