Fisheries Standard Review

Exploring alternative structures to the MSC Fisheries Standard and scoring approaches

Contributing to the MSC Efficiency Review Process

Consultant Report

November 2019

Inmara Ltd.

Marine Stewardship Council

Table of Contents

The views and opinions expressed in this report do not necessarily reflect the official policy or position of the Marine Stewardship Council. This is a working paper, it represents work in progress and is part of ongoing policy development. The language used in draft scoring requirements is intended to be illustrative only, and may undergo considerable refinement in later stages.

How to refence this report: Inmara Ltd. 2019. Exploring alternative structures to the MSC Fisheries Standard and scoring approaches, contributing to the MSC Efficiency Review Process. Fisheries Standard Review Consultant Report. Published by the Marine Stewardship Council [www.msc.org], [LINK], 36 pages.

1. Introduction

The report is the final project output of a consultancy project undertaken by Inmara Ltd. on behalf of the Marine Stewardship Council (MSC). Inmara Ltd were commissioned to explore potential alternative structures to the MSC Fisheries Standard and Scoring Approaches. The results of this project will contribute to the much larger process of MSC Efficiency Review as part of the MSC Fisheries Standard Review (FSR) process. The project was led by Mr Tristan Southall, of Inmara Ltd., with contributions and input from Dr. Rob Blyth-Skyrme, of Ichthys Marine Ecological Consulting Ltd. and Dr John Hambrey, of Hambrey Consulting.

1.1. Background to Project

1.1.1. The MSC Efficiency Review

MSC has stated that the overall objectives of the Efficiency Review are to reduce both the structural complexity of the MSC Fisheries Standard and scoring system complexity. Specific objectives are as follows:

- Identify any structural components of the Fisheries Standard, including both performance indicators (PI) and scoring issues (SI), that may not affect the outcome of assessments (i.e. correlated, invariant, or unlikely to trigger conditions).
- Identify areas of the Fisheries Standard that might be simplified while maintaining intent (i.e. with the same level of sustainability performance).
- Consider the application of alternative standard structures and scoring approaches, which would maintain the same level of sustainability performance but with reduced complexity.
- Develop proposed structural changes to the standard which, together with the move to a digital audit and assessment platform, will create an integrated, scalable system to aid in conducting assessments while improving data collection to support monitoring and evaluation across the program.

This project therefore comes at an important moment in the MSC's journey toward its goal of providing a simpler standard, which can be more readily understood by stakeholders and be reliably and consistently applied by assessors.

1.1.2. MSC Efficiency Review Workshop

A two-day workshop was held in April 2018 with 11 Conformity Assessment Body (CAB) and assessment team personnel, Assurance Services International (ASI) and MSC Fisheries Team members (including 2 members of this project team). The objectives of the workshop were in line with those described above. The key outcome from the workshop was a prioritisation of key issues to explore during the Efficiency Review:

1. Simplify and clarify language in the Fisheries Standard.
2. Further investigation and analysis of redundancy in Performance Indicators and Scoring Issues to determine if any Performance Indicators (PIs) or Scoring Issues (SIs) can be removed, without changing the intent or lowering the bar.
3. A re-structure of the Fisheries Standard logic:
   a. Integration of P1 and P2 management PIs and P3 fishery specific management PIs, thereby reducing the overlaps and gaps that exist in the current structure.

b. Restructure outcome, management and information PIs in P2 to include conditional scoring and a decision tree approach.

4. Aggregation of SI scores to determine Principle score, rather than aggregating SIs scores to determine the PI score and averaging PI scores to determine Principle scores.

1.2. Project Scope

This project builds upon the outcomes of the April 2018 Efficiency Workshop, described above, and seeks to explore these potential options. Although the outcomes described above cross all 3 MSC Principles, this project has placed most emphasis on Principle 2 as this was seen as the highest priority given the high level of existing complexity and inefficiency.

All proposed changes relate to the MSC standard and some of the proposed changes would also require changes to the MSC guidance. However, it is not anticipated that any of the proposed changes would require changes to the MSC certification process.

1.3. Approach & Methodology

This project is relatively quick and short. As such, it has not undertaken a comprehensive range of analysis or a comprehensive assessment of the impact of any proposed changes on the scoring of fisheries. Instead it considers the ideas raised at the April 2018 workshop and explores the potential merits of an alternative structure. However, the project does seek to provide a rationale for the proposed changes and give sufficient consideration of likely implications of those changes to determine whether further more in-depth analysis is justified.

In approaching this work, the core question, most frequently asked by the project team when seeking to clarify the existing scoring issues and scoring guideposts was:

“What is currently the exact requirement of the standard”?

In asking this, we sought to get an unambiguous understanding of both the intent of the scoring issue and the precise scoring thresholds. This is essential in order to understand what wording within current scoring guideposts are critical. For example, where an SI requires there to be “some quantitative information”, “adequate to assess the impact of the UoA” on a species “with respect to status” is this requiring just information on the level of UoA derived mortality (i.e. catch) or is it also asking that there is sufficient information to determine stock status? And how are words like “some” and “adequate” to be interpreted?

The proposed alternative P2 structure which is presented at the conclusion of this report, seeks to include the same scope of questions currently being scored within the standard and it seeks to maintain the same scoring thresholds, but it seeks to do so in a way which is clearer and therefore quicker to undertake. This is built on the foundations of the existing structure and is not intended to be a complete reinvention. The project team have sought to avoid any bias or influence to move the scoring thresholds, retaining a focus on how to improve efficiency without any reduction in robustness.

1.3.1. Initial Review

The initial review focussed on key areas of the existing MSC Fisheries Standard and the application of the existing standard in order to provide an evidential basis for the subsequent restructuring work. This focused on 2 key areas:

1. The degree of duplication and overlap (or possibly gaps) between scoring issues, or ambiguity in the language of the standard, initially within P2, but also between P2 and P3.
2. (Using the MSC Scoring and conditions database): The reason for low outcome scores and focus of resulting outcome conditions and how this relates to scoring of management and information PIs of the same component.

1.3.2. Internal Workshop

An internal workshop was held in London in January 2019. This involved all three members of the project team over two days. However, two members of the MSC Fisheries Standards team also attended. During this workshop the results of the initial review and potential solutions within a restructured standard were rigorously explored, allowing space for creative ideas.

When issues, gaps, or sources of potential confusion or inefficiency were identified, possible solutions were proposed, but these were always followed by consideration of whether the proposed solution was likely to reduce or add to complexity or efficiency. This provided an on-going reminder of the overall need for increased simplicity.

Proposed solutions or changes to the standard were ranked according to the scale of change required within the standard and guidance in order to be implemented.

Draft findings of this project were then presented and discussed in detail at the May 2019 Efficiency Workshop, in London. This provided an opportunity for both experienced MSC assessors and MSC staff involved in the project to comment. The final reporting outputs and the proposal for a revised structure of P2 reflect comments received.

1.3.3. Developing a proposed revised structure

In developing a proposal for a revised structure, we have explored how best to balance the need for simplicity with the requirement to ensure consistent application. Where there is complexity within the standard, we ask whether it is necessary. In all cases, the project team have sought to:

- Maintain the recognised structure of the 3 MSC Principles
- Maintain notional 60, 80, 100 scoring
- Maintaining the patterns of Conditions requiring action to improve to the SG80 level

1.3.4. Impact testing

No impact testing has yet been undertaken of the proposed changes. A process of impact testing would subsequently be required if the proposed structure changes are met with initial approval and are to be subject to further consultation prior to adoption. This impact testing should not only look at the impact of the proposed changes on scoring outcomes, but also at whether the proposed changes simplify and increase the efficiency of the process of scoring.

2. Findings of Initial Review

The initial review of the standard, in the areas relevant to the scope of this project, showed that there were a number of notable areas of challenge, that therefore offer the opportunity for improvement within the standard.

2.1. Duplication & Repetition

2.1.1. Repetition in justifications

There are instances of duplication in the evaluation table within P2, which are discussed in more detail below (2.1.2). These are the main focus of the Efficiency Review. However, when considering duplication, it should also be recognised that repetition is experienced by the assessor (and the report reader) when the fishery data collection processes or fishery management practices must be repeatedly described for many components in turn, or elements in turn, or UoAs in turn. For example, many aspects of a fishery information system (vessel monitoring, logbooks, observers etc), will be the same for primary species (whether main or minor), secondary species (whether main or minor) and ETP. But the way that the standard and the evaluation table are currently structured creates an artificial separation between elements and components, which does not recognise that the components and elements are essentially subject to the same management system[1].

2.1.2. Repetition in Scoring Issues

One of the factors which affects the time that it takes to undertake a fishery assessment is the amount of reporting time required to provide justification for each scoring issue. In order for the process to be as efficient as possible it is therefore important that there is no repetition, or more accurately, no overlap, between scoring issues. Very careful consideration should be given before adding further scoring issues at the time of any review of the standard, including consideration of whether the additional scope can be better included by a change elsewhere in the standard, such as within the definitions or guidance. This may achieve the same ends, but without additional time constraints being placed on the assessors or cost constraints being placed on the client fishery.

2.1.2.1. Within Principle 2

Perhaps the biggest cause of repetition within the scoring issues in Principle 2 is the question of whether or not management is working. This is clearly a very important question but it is currently asked in slightly different ways across many scoring issues, which all require a very similar level of scoring justification. Taking secondary species as an example:

- 2.2.1a: “(management) in place that are expected to ensure ……”
- 2.2.2a: “(management) in place that is expected to maintain ……”
- 2.2.2b: “(management) is likely to work ……”
- 2.2.2c: ”(management) is achieving its objective …….”
- 2.2.3c: “whether (management) is achieving its objective ……

Just one of these scoring issues (2.2.2b) is actually about evaluating the performance of management. This is the same for Primary species. There are a number of further examples of repetition within Principle 2. For example, an SI asks whether management is *in place* and a later SI asks whether management is *implemented*. It is not clear what the difference is between “in place” and “implemented” and it is likely that this will be interpreted in different ways by different

[1] The new requirement (at CRv2) for a review of ‘alternative measures’ also requires these to be described in 3 components when it is likely that there may just be a single such review of gear modifications (for example), which is applicable to 3 components.

assessors. Indeed, the Principle 2 Phrases table (SA8) defines “In Place” as “when a measure or strategy ….. has been implemented”.

Generally, there is less repetition within the information PIs in P2, with a clear separation between the need for information about impact and information to support management. The exception is the 5 SIs for ecosystem information which are likely to all refer to a very similar evidence base.

2.1.2.2. Across Principle 3 and the other Principles

The 2nd half of Principle 3 focusses on the Fishery Specific Management System and makes direct reference to Principle 1 and Principle 2. This is therefore the system which delivers the management described in P1 and P2. Unsurprisingly therefore, P3 sometimes describes within the assessment report, management which has already been described in P1 or P2. This repetition is also reflected in scoring and is often exacerbated because different Principles will typically be undertaken by different assessment team members, so each may be unaware, at the time of writing, what justification and description has been provided on the same topic in other Principles. Examples are provided below for each of the P3 fishery specific performance indicators.

> PI 3.2.1: Fishery Specific Objectives often refer to Maximum Sustainable Yield as the long term P1 objective and may even refer to the Harvest Control Rule ‘tool’ (i.e. catch limits), which was scored in SI 1.2.2c as the measurable short-term objective.
>
> PI 3.2.2: Decision-making Processes often describe the role of the Harvest Control Rule and stock assessment, which will have already been described in P1. As a result, the presence of a well-defined harvest control rule will result in an SG80 score for SI 1.2.2a but is also likely to contribute to an SG 80 score in SI3.2.2a.
>
> PI 3.2.3: Compliance and Enforcement scoring tends to focus, in many assessment reports, on P1 compliance, however the guidance makes clear that it should also include[2] consideration of compliance with MPAs and other spatial management measures. This therefore has considerable overlap SI 2.4.2d which relates to compliance with requirements to protect VMEs. SI 2.4.2d is the only place in P2 where compliance is specifically mentioned, even though compliance is a key factor in determining the efficacy of management across all P2 components.
>
> PI 3.2.4: Monitoring and Management Performance Evaluation will often describe the process of stock assessment or the review of the stock assessment, which has already been described in P1. Additionally, P2 has already asked about the level of evaluation across all components, so these reviews may again be referred to when assessing whether ‘all’ or ‘key’ parts of the management system are subject to review.

Although there is clear potential to more clearly differentiate between some scoring within P1, P2 and P3, this would likely require a much more significant restructuring of the standard and perhaps consideration of whether the existing division between 3 discreet Principles remains the most appropriate. As a result, it is concluded that this maybe too significant a change for this round of

[2] GSA4.9

standard review so has not been further prioritised at this stage. This does however fall under the 'Changes for the Future' discussed in report section 3.2.

2.1.3. A single issue impacting many scores

Because of overlap within the structure of Principle 2, a single issue or weakness within the management system may have knock on consequences across many scoring issues and even across many components. The approach that assessors take to multiple penalising of the same issue may differ. Many assessors state that it is not the intention to multiple penalise under different scoring issues for the same finding[3][4]. Other assessors may take a more literal approach and reduce scores across all scoring issues where the finding is relevant. For example, if there was a paucity of information about the catch profile, this could be seen as a single issue which could reasonably be addressed by a single condition and a single resulting client action. However, it is likely that this would currently affect scores across many areas of P2. For example, within a single component a lack of catch data could impact on scoring in the following areas:

- main outcome status (likelihood);
- minor outcome status (likelihood);
- management strategy evaluation (lack of UoA specific data);
- management strategy implementation (lack of UoA evidence that strategy is achieving objective).

Whereas the actual gap relates to a lack of information, so should ideally be addressed in the information component. The same gap relating to lack of catch data would have a duplicative impact on other P2 components (i.e. the same issue applies for both Primary and Secondary) and across all elements.

2.2. Multiple clauses within an SI / SG

In spite of having multiple scoring issues, there remain many cases where a single SI or Scoring Guidepost (SG) contains several clauses or questions. The inclusion of multiple clauses within an SI (or SG) makes the intent less clear due to a lack of plain English. This also means that when analysing a fishery performance (as part of MSC research into the positive global impacts which result from the programme), it may not be clear exactly what the cause of any score changes are. For example, if a fishery shows an improvement for a score against a particular SI, the justification needs to be scrutinised to see which clause within the SI (or SG) has led to the increase in score. This complicates and potentially undermines attempts to analyse MSC scoring patterns across fisheries or monitor performance improvements across the whole MSC program.

---

[3] This may derive from an understanding that the original assessment tree was based on the "expert choice" software, which applies Analytical Hierarchical Process (AHP) in a simple additive weighting method which is designed for scoring independent criteria.

[4] This is illustrated with reference to past TAB papers. During the development of the Fisheries Assessment Methodology the need for criterion to be independent in order to avoid unintended weighting or double penalising and for additive soring to make sense, was highlighted (TAB 11 (2007).

The clearest example of multiple clauses within an SI SG is for secondary species outcome at the 80 level, which states:

> Main secondary species are highly likely to be above biologically based limits
>
> OR
>
> If below biologically based limits, there is either evidence of recovery or a demonstrably effective partial strategy in place such that the UoA does not hinder recovery and rebuilding.
>
> AND
>
> Where catches of a main secondary species outside of biological limits are considerable, there is either evidence of recovery or a, demonstrably effective strategy in place between those MSC UoAs that also have considerable catches of the species, to ensure that they collectively do not hinder recovery and rebuilding.

This Scoring guidepost contains too many clauses variables to be readily understood (it includes 3 x "or", 2 x "and" and an "if") and therefore has the potential to cause considerable confusion.

2.2.1. Is outcome really outcome?

Following on from this, it is noted that the outcome status PIs in Principle 2 do not simply provide a scoring scale relative to outcome status. When moving from the SG60 to the SG100 the probability threshold also changes. This is deliberate as it reflects the need for greater precaution to be applied to scoring where information is lacking. However, as a result, low scores may be as a result of either information shortcomings or depleted status. Furthermore, if it cannot be concluded that the status is above PRI (or equivalent biologically based limits), then a further question is asked, within the same SI, to determine whether the management is likely to work. So, a low score for outcome status is actually never simply the result of poor status but is instead either the result of a lack of evidence, or the result of both a low status and a management gap (Figure 1). Although this may be clear if the scoring justification is read, it is not clear when any analysis of scoring is done.

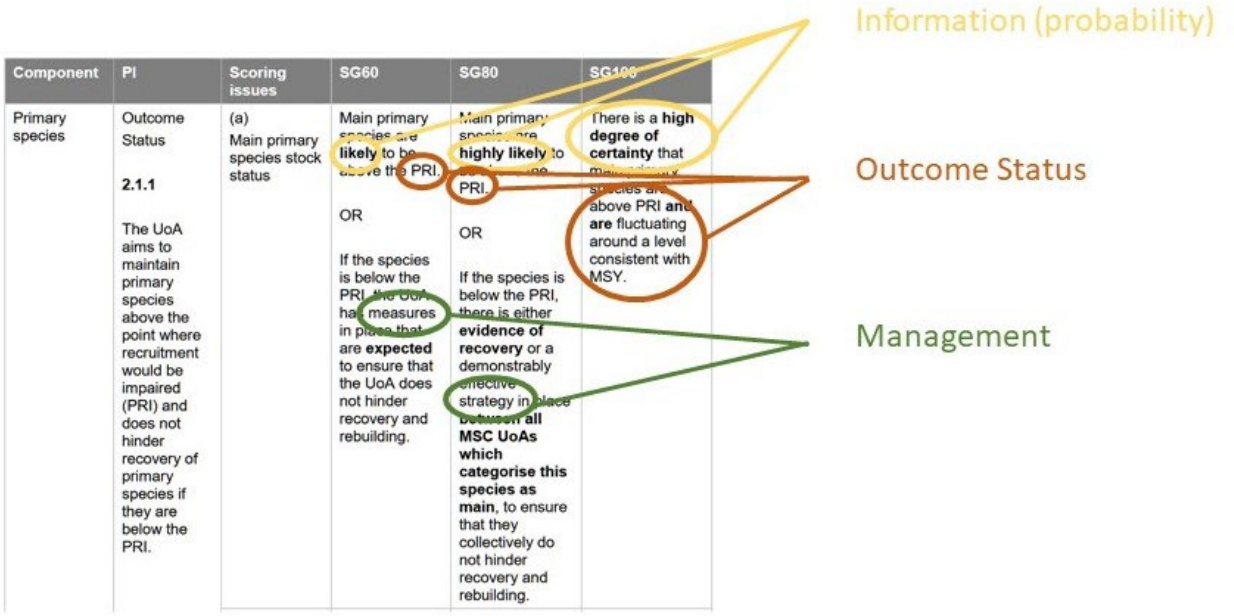

Figure 1: The outcome status scoring issue is heavily influenced by management and information

By contrast, when scoring outcome status with the RBF, the outcome score is simply calculated, so is a straightforward function of risk, with no sliding scale of increasing probability or consideration of management efficacy being applied at the same time.

The situation is even more pronounced when scoring both habitat and ecosystem outcome status as there is no change in the status threshold whatsoever at any of the scoring guideposts. Instead higher scores are only achieved by higher levels of confidence or probability, which is likely to be the result of improved information (or higher levels of assessor confidence), rather an actual improvement in status. So, it is arguable whether the current score for habitat or ecosystem outcome status is actually scoring outcome status.

With the exception of Primary species or scores derived using the RBF there are typically no quantitative reference points to support scoring of outcome status in the many complex areas of the natural ecosystem which P2 seeks to score[5]. This means that outcome scores are inevitably heavily reliant on the level of expert confidence, rather than a quantitative probability assessment. Scores given in this way are often those which are most likely to be challenged by other experts (i.e. stakeholders). Although the guidance introduces thresholds of probability (for example when considering the likelihood of reversibility of habitat impact) because there is typically no probabilistic measurement of impact, the quoted thresholds of probability are rarely used so only create the impression of quantitative scoring, when in fact the process is non quantitative.

The clearest example of this relates to habitat outcome status. In order for a score of 80 to be achieved it must be demonstrated that: "The UoA is highly unlikely to reduce habitat structure and function to the point where there would be serious or irreversible harm". When the guidance is reviewed what this literally means is: *There is a less than the 30th %ile probability that the UoA impacts to the point where the habitat (both within and beyond the managed area) would be unable to recover to 80% of the structure and function to which it would eventually recover to (within existing environmental and anthropomorphic conditions), within 5-20 years if fishing were to cease entirely.*

By contrast a score of 60 would require a probability below the 40th %ile and a score of 100 would require a probability below the 20th percentile. So, the difference between a score of 80 and a score of 100 is entirely due to a change in probability of an 80% recovery within a range of years. In reality very few assessments make a fully quantitative justification for the score given and more often expert judgement is the key determinant of the final score. It is perhaps therefore of no surprise that the scoring of this PI is often the source of stakeholder challenge, either at the stakeholder review stage or at the objections stage of the assessment process.

2.2.2. Outcome conditions

As a result of the conflation of information and management efficacy within the outcome status scores, the resulting conditions on the outcome status PIs are likely to require action to either improve information, or improve efficacy of management or even demonstrate that the element in question is not "main". This is an example of where low scores in one PI (outcome status) are

[5] This issue of the lack of quantitative benchmarks for many P2 components was highlighted in the past at TAB meetings to review the development of the Fisheries assessment methodology (TAB 12 2007).

actually caused by shortcomings in another PI (information). For example, where an improvement in the outcome status score occurs this may be the result of either: (i) improved outcome status; (ii) improved information leading to increased confidence; or (iii) more effective management meaning the UoA is unlikely to hinder recovery.

The following are examples of conditions triggered by scores of less than 80 in outcome status PIs which require action in relation to information or management. All are derived from currently certified fisheries:

- VME habitat status. "While there is evidence that it is unlikely that derelict FADs reduce structure and function of the VME habitats to a point where there would be serious or irreversible harm, due to the potential impact over a number of years and lack understanding of the real nature of the issue, it cannot be concluded that this is highly unlikely. More evidence is required".
- ETP status: "Although most of the scoring elements scored at SG80 here, scoring element Sebastes norvegicus did not. The client recorded both redfish species together in the bycatch, rather than separating the two species out, when determining weight and percentage bycatch. The accuracy of the catch information on Sebastes norvegicus needs to be improved through better recording to species level.
- Habitat outcome status: "There is insufficient evidence to be able to state that the fishery is 'highly unlikely' to reduce habitat structure and function of soft-bottom sponge communities, hardbottom coral gardens and sea pen fields to a point where there would be serious or irreversible harm. All vessels in the UoC need to provide data on i) any known interactions with VMEs (e.g. any relevant benthic species attached to the trawl) and ii) the location of fishing activities, so that interactions can be evaluated with more precision and certainty.

Many more examples, such as those noted above are evident in the conditions database. Having reviewed the database of existing P2 conditions, it is also clear that in many cases, where the requirements of a condition on outcome status are for improved information or management, there is a further, duplicative condition on the management of information PI, or in some cases a combined condition.

As a result, a large number of conditions, including many that fall under the outcome PIs, require improved evidence. Whilst these have generally resulted in significant efforts by fishers to comply, the quality of information collected and the level of resulting analysis, even where well-resourced and planned, may be insufficient to lead to a change in the outcome status score. For example, the additional information could result in greater evidence of poor status, meaning that the client fishery's actions failed to lead to an increase in score.

2.3. Dividing the catch

The MSC requires that all species which may be impacted by the fishery are properly considered as scoring 'elements'. Different species are scored against slightly different criteria. This makes sense because the expectations in terms of information and management will be different depending on the proportion of the catch and whether the species is actively managed or protected. However, the MSC standard currently requires the catch to be divided into a very high number of different groups, some of which are obvious and easily understood, but others of which are more esoteric or arbitrary and less easily understood by stakeholders.

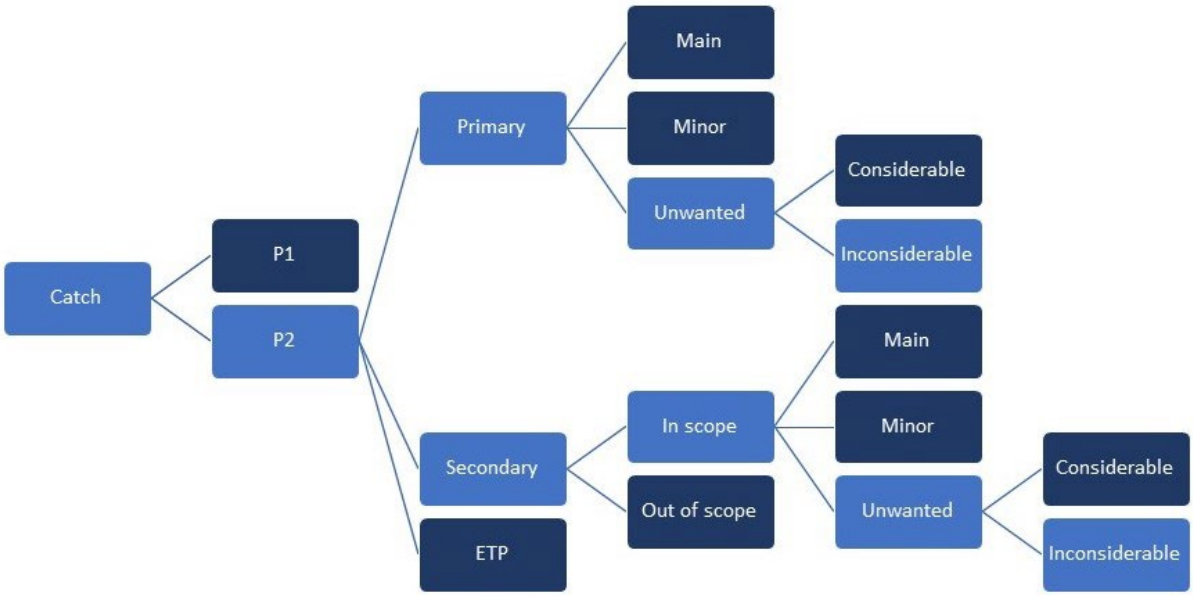

Figure 2: The different categories that the catch elements must be divided into (Dark blue represents the point at which an element is scored. Light blue indicates that a further categorisation is required).

This project has considered whether the current groupings into which the catch is divided are necessary and justified in order to reliably score the fishery or whether different (and fewer) groupings might enable more efficient, but equally robust, assessment.

2.3.1. Into components

The catch is initially divided into components. The separation out of ETP species for consideration allows for a higher bar in terms of management and information requirements to be applied. Although the definition of ETP can be complex, due to differing legislation and inconsistencies between jurisdictions, it is nonetheless tightly defined and readily understandable to stakeholders[6]. Similarly, it is appropriate to give habitat species separate consideration as characteristics of management and data collection will differ. However, the separation of the remaining catch into Primary and Secondary is less obvious and requires a good level of understanding of the MSC standard.

2.3.1.1. Primary & Secondary

As very few management systems will recognise the difference between primary and secondary species, it is likely that explanation will be required to enable stakeholders to understand this categorisation. The fisheries information systems and, in many cases, the management systems which pertain to the fish species within the catch are likely to be the same, regardless of whether there are reference points in place[7]. As a result, there is often considerable repetition in the scoring justification for Primary and Secondary components.

---

[6] It is noted that there is a separate MSC project currently on-going as part of the FSR, examining the definition of ETP species.

[7] Reference Points are the key determinant of whether a species is classified as Primary or Secondary.

In previous versions of the MSC standard (i.e. FAM to CR v1.3) the first two components of P2 were defined as 'Retained' (i.e. those species that were typically landed to market) and 'Bycatch' (i.e. those species which were typically discarded). When these definitions were changed, there was perhaps a preference to retain five P2 components and to separate out those species which can be simply scored using reference points and which are most likely to be the next candidates for a future P1 assessment.

There are also some minor differences in the scoring bar between primary and secondary. For example, the scoring of likely, highly likely and high degree of certainty have a slightly lower percentage threshold for primary than for secondary. However, it is unclear what the practical application of this lower threshold is, given that the secondary score will most likely be determined by RBF (in which case the probability is not scored) and if not, it is unlikely that there would be a probabilistic stock assessment with reference points (because if available the species would likely be considered primary).

The other differences in scoring relates to the threshold catch percentage of any species below PRI (for Primary) or biologically based limits (for secondary) which triggers a requirement to consider cumulative catches of other MSC UoAs. For primary the threshold is as per "main" (i.e. 5% or 2% if less resilient) for secondary threshold is 10% (defined as 'considerable')[8].

2.3.1.2. Unintended Scoring consequences

There are some unintended scoring consequences to this division of components. If a fishery has zero impact on a component then it receives an automatic score of 100 for outcome status[9]. This rewarding of the fishery makes sense where there is zero impact on other species within the ecosystem, so for example, the most selective fishing gears receive the highest scores. This also made sense in previous versions of the standard to reward fisheries with no other retained catch or no bycatch. However, with the current definition of 'Primary' and 'Secondary' it now rewards a fishery which occurs within a jurisdiction where no species meet the MSC definition of Primary. For example, if a management authority does not yet manage any species relative to reference points (or where the species being assessed at P1 is the first species within the jurisdiction to be managed according to reference points), then there will be no species designated as Primary. In this situation, even a non-selective fishery with a high bycatch will score an automatic 100 for Primary species. This creates an artificial and potentially significant increase in the principle level score. By contrast, a similar fishery within a neighbouring jurisdiction which have taken initial steps to begin management of some commercially important species by reference points, runs the risk of getting considerably lower scores because the Primary Species PI must be fully scored. This has the potential to reward lack of management and as such poses a reputation risk to the MSC.

2.3.2. Into "Main" & "Minor"

Primary and Secondary species must then be further divided into "Main" and "Minor". Historically, some assessments only referred to "main" species, so it has been beneficial for consistency to

---

[8] This 10% threshold for consideration of cumulative impacts also applies to non-ETP out of scope species, even though a negligible catch of these species would trigger a classification of "main". So the cumulative impacts of MSC fisheries on a "main / out of scope" species would not be considered unless more than 10% of the catch weight in each fishery.

[9] SA3.2.1

tighten up the definition of exactly when a species should be considered and when it shouldn't. However, the current mechanism for dividing between "main" and "minor" and then scoring is complex and time consuming. Although guidance is provided on the percentage thresholds for "main" or "minor", these are not fixed, so require a further consideration of vulnerability or resilience and the total catch of the UoA relative to other fisheries to determine whether the default percentage thresholds should be overridden. The minimum percentage threshold for minor species is not defined. Therefore, in a mixed fishery the assessor should in theory undertake scoring on all species however negligible a contribution they make to the catch, in order to demonstrate consideration of "all" minor species.

2.3.2.1. Unintended scoring consequences

As minor species only affect scoring in the 80 to 100 range, there is the potential to only score minor species when necessary to provide a positive boost to scores. Minor species will generally not have a negative influence on scores. Conversely, where a species is on the cusp of being considered "main", this determination may have a significant influence on scores. The separation of species into "main" and "minor" which are addressed in different SIs creates additional complexities in terms of scoring, as evidenced by assessors seeking clarification of how to score in circumstances where there are no main species but minor species are depleted[10].

2.3.2.2. Out of Scope "Main"

Out of Scope species are those which are reptile, mammal, amphibian or bird. However, these will be defined as ETP and scored separately where they have an IUCN status of Vulnerable, Endangered or Critically Endangered (or where they are protected by applicable national legislation or international conventions). Any Out of Scope species which are not ETP, must automatically be considered "Secondary main" and therefore fully scored across all secondary species PIs (unless released alive with a high potential for post-capture survival). So, if there is mortality of a single bird (which is not vulnerable or endangered or protected) it must be fully scored at the SG60 and SG80 levels for outcome, management and information, meaning that it must be subject to a partial strategy supported by some quantitative information. Given the lack of vulnerability or protection in place for the bird (because it is not ETP), it is likely that the level of management and information may be low and status will be undefined so will require the use of the RBF. Because there is no potential to apply a filter based on the scale of risk or negligible frequency of occurrence in the catch, this places a huge potential time and cost implication on the assessment. And the result will often be either a poorly justified 80 score or a condition which may be seen by stakeholders as requiring resources to be targeted to a low risk issue.

2.3.3. Into Elements

Finally, once the catch has been divided into all the many groups described above, each species within that group must be scored separately as an "element" which contributes to the overall score. So, if there are several species, all with a similar catch percentage, the same management, the same level of information and the same status they must all be scored separately. Technically this

[10] https://mscportal.force.com/interpret/s/article/P2-species-outcome-PIs-scoring-when-no-main-or-no-minor-or-both-PI-2-1-1-1527262009344

separation should continue to apply across all of the scoring issues within the 3 PIs of the component. This creates huge workload and huge repetition[11].

In order to combine the scores of many elements into a single score table 4 of the FCR is applied using the rationale of whether 'few', 'some', 'most' or 'all' elements meet a particular scoring threshold. This is based upon a sound logic to ensure that where a single element fails to meet either 60 or 80 the fishery will either fail, or trigger a condition, regardless of the high scores for other elements. However, it is nonetheless a further source of complexity and confusion which can be difficult to explain to stakeholders.

At present when undertaking the RBF, it is possible to group species together, where their characteristics and resulting scores can be shown to be similar. However, no such grouping is permitted within the standard evaluation table, so all elements must be described in turn.

The number of elements that are now assessed is large, and the cost of comprehensive assessment extremely high. This disadvantages small scale fisheries which are often ironically less environmentally damaging.

2.4. Scoring

The MSC scoring process has gone through several evolutions. Issues such as scoring thresholds, partial scoring, rounding-up of scores and combining multiple scores have all been addressed and clarified. As a result, the scoring process is probably more consistently applied. However, the process is complex. In order to determine a Principle-level score the Performance Indicator scores are numerically calculated (an average for P2 and a weighted average for P1 and P3). However, the PI score is not numerically calculated and is instead determined by combining the SI scores using the "few", "some" and "most" logic described in FCR 7.10.5. Where multiple elements contribute to the SI score this is determined with reference to table 4 in the FCR. Because different PIs have differing number of SIs the scoring possibilities change:

- A PI with only 2 SIs can only score 60, 70, 80, 90 or 100 (assuming there is only 1 element)
- A PI with 3 or 5 SIs can be scored 60, 65, 75, 80, 85, 95 or 100 (again assuming only 1 element).
- Where there are 5 SIs, if 3 are met at SG100 and 2 are met at SG80, the score is 95. But if a 4th SI also meets SG100, meaning that only 1 was scored at SG80, the score remains unchanged at 95.

To add further complexity, some PIs have differing number of scoring guideposts across the SIs, so the scoring calculation changes, depending on whether scoring is occurring at the SG60 level, the SG80 level or the SG100 level. This system is understandable to qualified assessors (once used to it), but the question should sensibly be asked (as part of the standard and efficiency review process) whether this is really the best and clearest system possible and whether the benefits of the original

[11] It is interesting to note the evolution of 'elemental' scoring within the MSC to the point where now every species must be considered. At the time of TAB 12 in 2007, which focussed on the development of the Fisheries Assessment Methodology, it was noted that it would be unrealistic to expect every species of bycatch or ETP to be scored individually.

logic are outweighed by the confusion and likelihood of scoring errors caused by the level of complexity.

2.4.1. The 100-point range

During the April 2018 efficiency workshop, consideration was given to whether the 100 point scoring range was a useful characteristic of the MSC standard. Legitimate questions have been asked about scoring detail below 60 (for example for fisheries scored at less than 60 wishing to show progress during a FIP). However, scores below 60 are not defined. At the other end of the scoring range it has been asked if scores greater than 100 can be achieved. The scores are sometimes thought of as percentages which they are not.

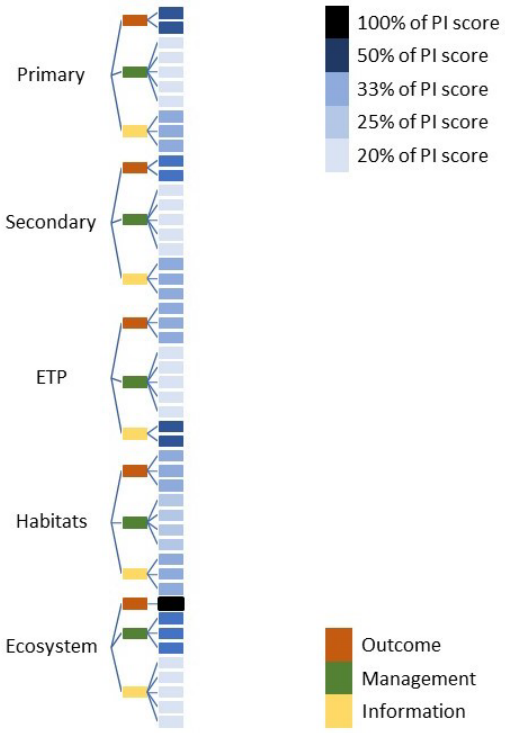

Figure 3: The current percentage contribution of SIs to the overall P2 principle level score.

An alternative scoring model is used in other audit programmes simply based on the number of minor or major non-conformities or a maximum permitted number of conditions. Whilst there was some enthusiasm for adopting this simpler scoring system, especially one with no requirement for scores to be calculated, it was concluded that over time the understanding of an MSC 60 being the minimal acceptable level, an MSC 80 being Best Practice and MSC 100 being "State of the Art" have become well understood across the industry. Many fisheries across the world were working to improve scores within this overall range. So, it was concluded, at this stage, that the scoring range of 60-100 should be maintained (even if the method of score calculation changes).

2.4.2. Weighting

One of the unintended consequences of the existing MSC structure is that because different PIs have different numbers of SIs, each SI makes a differing contribution to the overall principle level score. Within P2 an SI, which is the sole SI within a PI will contribute up to 6.7% of the overall Principle level score (1/15th of the overall score) whereas an SI which is one of 5 SIs contributing to a single PI will only contribute up to 1.3% of the overall Principle level score (Figure 3). At the extreme, in P3 the single scoring issue for Long Term Objectives (3.1.3) contributes up to an enormous 17% of the overall Principle level score[12].

The structure currently allows for weighting to be applied at the component or PI level but not at the SI level. Within P1 this allows for slightly more weight to be applied to outcome status, but in P2 the weighting between components is now equal, meaning that PI scores are simply averaged across the Principle. However, over the years more emphasis has been placed on SI level evaluation and justification, with conditions now focussed at the SI level. At the same time, more SIs have been

[12] This PI also has a very small difference between the scoring guideposts at SG80 and SG100, so a very subtle interpretation can have a very big influence on scores.

added to some PIs without consideration of the fact that this reduces the weighting of other SIs within the PI.

During the early evolution of the MSC standard the weighting which was applied was transparent and intentional. Careful consideration should be given to whether the weighting that is currently applied is both necessary and intentional.

If there is a need for intentional weighting then this should be presented in a transparent manner.

2.5. Summary of Initial Review

Given the outcome of the April 2018 efficiency workshop and the findings of the initial review undertaken within this project there appears to be a strong rationale for change. The drive for efficiency and simplicity is not only about potential time and cost savings; it should also bring improvements in clarity and accessibility of the standard and increased consistency and transparency in scoring.

The highest priority for change appears to be in P2 as this has the potential to achieve greatest efficiency improvements.

3. Proposed Changes to Structure

Many potential areas of improvement have been identified in this study, some of which are simple and some of which would require a more wholesale change. In proposing a revised structure, we aim to have a proposal which has a reasonable expectation of being considered for inclusion in the next revision of the MSC standard (following the appropriate impact testing and the conclusion of the standard review process). In order to achieve this, the proposal aims to remain within the following constraints:

- No change to overall structure of 3 core MSC principles.
- No change to the intent of scoring thresholds • No change to the concept of scoring by SI, with guideposts at 60, 80 and 100
- No change in the requirement for an overall pass mark of 80 at the principle level;
- No change in the requirement for conditions to bring 60 level scores up to 80 level.
- No new acronyms or terminology[13] (although some terminology maybe removed).

3.1. A Hierarchy of proposed changes

We present suite of potential changes, which have been discussed both at the Efficiency Workshop in 2018 and this project's internal workshop in January 2019. These changes range from the small to the significant. We have therefore sequenced the proposed changes below in order of perceived priority or scale of change.

- Clarify Language and reduce duplication
- More tightly define "Measures", "Partial Strategy" and "Strategy".
- Simplify Grouping of P2 elements (remove "main" and "minor")
- Re-order Information, Management and Outcome
- Simplify scoring
- Reduce number of P2 components

The initial proposed changes require minimal change to structure and would be easily adoptable. Those proposed changes further down the list may require some structural changes and changes to MSC guidance.

[13] If Primary and Secondary species are combined then a term will be required for this new category of "other" species within the catch.

A revised P2 scoring assessment tree, which reflects all of these proposed changes is contained in Appendix 1 of this report.

Although there were more radical proposals for change discussed at both the April 2018 Efficiency workshop and in the January 2019 internal project workshop (such as consideration of whether the division of fisheries into 3 discreet Principles is realistic), we do not propose these here for consideration in this round of MSC Standard Review. These ideas for the future, which may have considerable merit with further investigation, are discussed in more general terms in section 3.2.

3.1.1. Clarify Language and reduce duplication

Perhaps the most important step in the MSC Review Process, is to carefully go through all scoring guideposts and check that the intent is clear and that the scoring thresholds are unambiguous. In doing so it should also be asked whether this scoring guidepost could be expressed more simply, in plain English. This is especially important given the high number of MSC practitioners (whether on the assessment side, the client side or the FIP side) with English as a second language[14].

3.1.1.1. Scoring Guideposts should just address SI

For example:

- P2 outcome status SIs should not ask about the quality of information or effectiveness of management.
- Management SIs should not separately ask whether management is 'in place' and implemented'.
- Management SIs which ask whether management is in place, should not also ask whether it is achieving it's objective (which has already been asked under the 'evaluation' SI).
- PI2.5.3c: Asks that the main functions of the components are "known" at the 80 level and "understood" at the 100 level. It is not clear what the difference is between know and understood.
- PI3.2.2b: Focusses on the responsiveness of the decision-making process but introduces a requirement for transparency at the SG60 level, which contradicts PI3.2.2c which is focussed on transparency.
- PI3.2.2e: Focusses on Approach to disputes, but asks 3 different questions at the 3 scoring guideposts.
- PI3.2.4: Assessors take a very different approach to this PI, in particular over whether a holistic evaluation of the overall performance of the fishery management system should be addressed in SIa or SIb. The scoring guideposts for SIb indicate that this is focussed on the fishery specific management system, but the title of the SI implies that the focus is on the degree of internal or external review.

3.1.1.2. Minimise the requirement cross-reference the guidance

[14] Review of past TAB papers also shows that this issue has been highlighted a number of times before.

It should be recognised that frequent cross-referencing to the Guidance adds time and is a source of inefficiency. There are instances where the scoring guidepost could be re-written to avoid this step. For example:

- Where the guidance provides a % threshold of likeliness, this % could be included in the scoring guidepost.
- Within P2, avoid having different percentage probabilities for 'likely', highly likely' and 'high degree of certainty' for different components. This level of complexity is not justified given the frequent lack of data presented in a probabilistic way.
- Where there is a requirement to consider "considerable" catches, this % threshold could be included in the guidepost.
- Where the guidance provides definition for the frequency of terms such as regular and biennial, this could simply be included in the guidepost.
- Where the guidance defines terms such as "serious and irreversible" consideration should be given to defining this with the scoring guidepost.

3.1.1.3. Remove duplication

The best way to remove duplication is to ensure that the scoring guideposts thresholds are only related to the specific focus of the scoring issue. A further way to also dramatically reduce repetition would be to reduce the number of arbitrary categories that the catch must be divided into. This is discussed in section 3.1.3.

3.1.2. More tightly define 'Measures', 'Partial Strategy' and Strategy'

Considerable time is spent in assessments on debating whether management can be defined as "measures", "partial strategy" and "strategy". Furthermore, additional description is often contained in the scoring guideposts which adds to and sometimes confuses the description in the guidance. Scoring would be clearer and less ambiguous if the terms "measures", "partial strategy" and "strategy" were much more tightly defined with only a single definition in use throughout the standard. Given the many factors that contribute to the determination of the management level, this description is perhaps best left within the guidance, rather than in the scoring guideposts. In the table below we present what we understand to be the existing distinction between the 3 terms, based on both what is defined in the guidance and what appears in the scoring guideposts. We do not propose to change the definition, merely clarify it.

	Measures	**Partial Strategy**	**Strategy**[15]
Scope	UoA or wider		UoA *and* wider
Objective	Limiting impact / not hindering recovery SG60 Outcome status	Limiting impact / not hindering recovery SG80 Outcome status	Defined Management target SG100 outcome status

[15] For ETP there is an additional level of "Comprehensive strategy" however the only distinction with "strategy" is that it is "complete and tested". Although reference is made to monitoring, analysis and responsiveness, these appear to be required at the "strategy level".

Design	Either designed for component, or incidental		Designed for component
Linkages	Unlinked	Some cohesive links	Strategically linked
Responsiveness	Non-responsive	Response where shown to be ineffective	Fully responsive
Cumulative	Just UoA	UoA and other MSC fisheries	All fisheries
Direct / indirect impacts	Direct only		Direct & Indirect
Monitoring		Some	Full

Guidance to PI 2.1.2a and 2.2.2a says that the management arrangements (measures, partial strategy and strategy) relate to the management undertaken by the UoA. This appears to be the case for both 2.1.2a and 2.2.2a as the SG60 and SG80 require that measures / partial strategy are in place for the UoA to ensure that it does not hinder recovery. By contrast, there is no reference to the UoA in the SG100 and furthermore it refers to a strategy "to manage" the species (as opposed to avoiding hindering recovery). This definition of strategy implicitly includes a requirement to go beyond the UoA to enable management of all sources of mortality. This also ties in with the SG100 information requirement for monitoring, which some components refer to.

3.1.2.1. Move "Alternative Measures" to the management definition

A good management policy should keep abreast of developments of alternative measures. The degree to which a management body considers alternative measures is an indicator of the level of management. It could therefore be possible to include the level of review of alternative measures within the definition of the management level. Indeed, it is noted in the guidance (GSA 314.2) that 'a strategy should include regular review of alternative measures.

	Measures	**Partial Strategy**	**Strategy**
Objective	Some review	5 yearly review	2 yearly review

The 'Alternative Measures' could therefore be removed from a dedicated SI within the standard and included within the definition of management level described in the guidance. In order to meet the definition of 'measures', 'partial strategy' or 'strategy' all requirements would need to be met.

3.1.2.2. Shark Finning

Although shark finning SIs are being addressed in another FSR project it would seem more sensible for the shark finning requirements to be included within the guidance or scope. These should still be normative, but need not be in a dedicated SI which creates repetition and requires assessors to provide justification even when sharks are absent.

3.1.3. Simplify grouping of P2 species

The separation of "main" and "minor" is arbitrary, adds inefficiency, complexity and repetition into the assessment process and can have unintended scoring consequences. It is also likely to be a source of scoring error. Addressing this issue could create greater simplicity with no loss of rigour (as elements would still be properly considered). It would be preferable to simply state that:

- All species comprising more that x% of the catch should be scored as elements.

In order to maintain the current level of rigour within the standard, the cut-off threshold for inclusion should be less than the current "main" percentage, but should remove requirement to assess species caught in negligible quantities. A figure of 2% was suggested during discussions for this project, but the exact figure could be the subject for further impact testing. The more caveats and exceptions to this threshold contained within the standard, the greater the potential for confusion and inefficiency, so these should only be included where clearly required. One such case could be for 'Out of Scope Species', where a reduced threshold could be applied – for example 1% of the catch (bearing in mind that, by definition, these species cannot be rated by IUCN as critically endangered, endangered or vulnerable). In both these cases, catches below the specified percentage would be considered negligible and should not be scored[16]. The negligible elements of the catch, could still be listed within the report, to ensure transparency.

Having a different percentage again for the determination of "considerable" catches, which require consideration of cumulative impacts should also be reconsidered. The requirement for consideration of cumulative impacts could be included within the definition used to determine the management level.

Currently, the requirement to consider catches in other MSC UoAs in certain circumstances has the potential to create a large additional work-load which may be difficult to predict at the point of tendering. The higher the number of MSC fisheries, the greater the time requirement. This requires that assessors review the catch percentages in potentially numerous other assessments. Each time a new fishery becomes certified the previously undertaken analysis would need to be updated by all fisheries.

3.1.3.1. Must every element / species be described and scored at every SI?

Further consideration should also be given to which SIs require to be scored by elements. This should be more explicit within the standard. For example, when describing information for ETP species, should this exercise be undertaken for every ETP species within the managed area, or could the overall level of information pertaining generally to ETP species be scored. Or should it be somewhere in between where species are described and scored by groupings. The ability to group elements within justification and scoring, where similarities exist, should also be given greater consideration as this could also lead to increased efficiency.

Finally, for those SIs where it is determined that elemental scoring is necessary, consideration should be given to how to improve the reporting template to enable this. The current lay-out of the

[16] As "main" and "minor" do not apply to ETP species, so this percentage cut-off threshold of catch would not be applied to ETP species. ETP would be scored where present in the area of the fishery with the potential to interact.

assessment tree is not well suited to scoring multiple elements and clearly presenting how the overall score for the SI is determined. This is currently likely to be the source of scoring error. In considering a new lay-out for the assessment tree, the potential for automatically determining an overall SI score from multiple elements should also be explored.

3.1.4. Re-order Information, Management and Outcome

Information, Management and Outcome are linked and co-dependent. With poor information, good management should be more precautionary. Outcome status will be less well known, but with sufficient management precaution there can be increasing confidence. Indeed, outcome status will only be well known where management has required the collection of information to assess status. An understanding of outcome status is therefore the result of the management and information processes. And the outcome status is an indicator of management effectiveness or, it could be argued, an indicator of the necessity of management.

A possible approach, which was considered but ultimately rejected, whereby outcome status scores are capped, based on information and management scores.

		Management Score		
		60	80	100
Information Score	60	cap @ 60	cap @ 70	cap @ 80
	80	cap @ 70	cap @ 80	No cap
	100	cap @ 80	No cap	No cap

It would therefore make sense to begin an evaluation by assessing the level of information. This will inform the level of management, so it makes sense for management to be scored second. Finally, the outcome status is the result of management and informed by information, so it makes sense for this to be scored third.

This creates a further time saving and reduced repetition because there would no longer be any requirement to discuss either the quality of the information or the effectiveness of management when initially scoring outcome status. The scoring would proceed in a more logical and sequential manner.

3.1.4.1. Group by “Information” and “Management” rather than by component.

In the current P2 structure all PIs relating to a component are scored in turn (i.e. outcome status, followed by management, followed by information) before moving on to the next component. The proposed change of structure scores the information SIs for each component in turn, before moving on to score the management SIs for each component. The advantage of scoring information across all components initially is that it should present a clearer picture of the information processes and make it easier for the reader where an information process applies to more than 1 component. This also focusses the assessor on the overall adequacy of information across the fishery. Within the current structure, the information process is described for 1 component but the scoring moves onto the outcome status and management before again returning to describing a very similar information process. This creates duplication in the justification, inefficiency and a lack of logical flow in scoring justifications.

3.1.4.2. Should outcome score be constrained by information and management?

If there is poor information and poor management (i.e. management has not applied a level of precaution reflecting the paucity of information) then can the outcome status ever be concluded to be good? Should the adequacy of information be directly linked to (and effectively cap) the conclusion of the assessment of outcome status. If outcome status is concluded to be good and information is poor, then there will always be a risk that the collection of more information, will lead to a downgrading of the assessment of outcome status. This would effectively penalise a fishery for undertaking data collection.

During this project we therefore explored the potential of placing a cap on outcome status score linked to the scores achieved for management and information. Indeed, this was one of the original ideas behind the restructuring proposal. Various possible scoring combinations were considered. However, it was ultimately concluded that there were many exceptions that could be pointed to in cases where expert judgement concluded that the level of interaction or impact was likely to be negligible but because of the perceived low risk little information or management was available[17]. In this project we have concluded that a rigid cap on outcome scores would represent additional complexity and constrain the assessor's ability to make sensible determinations on a case by case basis.

3.1.4.3. Recognise Outcome Status as an indicator of Management Effectiveness

Outcome status will always be an indicator of management efficacy, so this could be included within the scoring of management as the outcome of the management evaluation. This would mean that outcome status is still scored (with the scoring guidepost being more tightly focussed on outcome status rather that likelihood or management), but that it scored at the end of the management PI after all information and management has been considered.

An alternative proposal was raised by assessors at the May 2019 efficiency workshop, which stated that outcome status informs the level of management, so it may be logical to present outcome status after information, but before management. However, it was noted that the "if necessary" caveat that applies to the requirement for measures or a partial strategy for all P2 components (except ETP), is determined based on the absence of interaction, not outcome status. Therefore, where there is an interaction, management will always be necessary, regardless of outcome status. A positive outcome status does not currently alter the need for management.

3.1.5. Simplify scoring

As every SI must be scored it would be quicker and easier if resulting SI scores were combined in a one-step calculation to determine the overall Principle Level score. As conditions now apply at the SI level (i.e. focussed on a particular scoring guidepost), there is no benefit to combining SI scores into a PI score, which in turn contributes to the overall principle level score. This would reduce a layer of complexity. As such this would greatly simplify the process of developing stakeholder understanding of the standard. This would also speed up the scoring process and reduce the likelihood of scoring

[17] In most cases these examples were for negligible quantities of "out of scope" species which are currently scored as "main". If this is changed, then it may be worth looking again at constraining outcome scores depending on the quality of information and management.

errors. This would also greatly simplify any future transition of the assessment scoring onto an on-line platform.

If it was decided that a particular SI should have a greater contribution to the overall principle level score (for example the single SI within the proposed structure on outcome status) then a weighting could be applied at the SI level as part of the single calculation. This would remove many of the complexities of current scoring as described in FCR 7.10.

3.1.6. Reduce the number of P2 components

We are aware that other projects are being undertaken within the MSC standard review process which are looking at some of the other components within P2. It is not the intention of this project to influence or prejudge the outcomes of these projects, however, it is difficult to properly cover the intended scope of this project (i.e. the goal of increased efficiency and clarity coupled with reduced complexity and repetition) without giving some consideration of the utility of the current P2 component structure.

3.1.6.1. Primary and Secondary

Careful consideration should be given to whether the value of splitting species according to the definitions of Primary and Secondary is sufficient to warrant the considerable complexity, inefficiency, duplication and unintended scoring consequences that result from this artificial division. Although some species have stock assessments and reference points and can be scored against PRI (i.e. Primary under the current definition), whereas for others (i.e. secondary) the RBF is used to determine outcome status score relative to “biologically based limits”, it would still be possible to score these under the elemental scoring approach without needing to first separate them into Primary and Secondary components. All but the most negligible elements are currently scored and this would still be the case even if the component structure was revisited as suggested. Although there are some differences in the scoring thresholds and wording between Primary and Secondary these are relatively minor, so the intent would not change significantly. These could be grouped together as “Other species”, “Other catch” or “Sundry species”, or some title with similar intent. This would still include any non-ETP out of scope species (albeit with a minimum catch composition cut-off applied). For clarity these could be scored after both ETP and habitats and would then encompass all other caught species.

The recommendation from this project would be to merge Primary and Secondary into a single component; “Other Caught Species” and for this to be scored after ETP and habitats, but other options could also be considered.

- Score commercial and non-commercial species as separate components (this is similar in intent to the old “Retained” and “Bycatch” components).
- Score “in scope” and “out of scope” as separate components
- Retain Primary and secondary, but with the simplified structure for these as proposed for other components.

3.1.6.2. Ecosystem

Careful consideration should also be given to whether the nine SIs within the ecosystem PIs add value to the existing structure. Clearly there it is important to consider ecosystem information and ecosystem management but these ecosystem considerations are already considered to some extent

in P1 (in relation to consideration for the ecological role of the stock) and in the other P2 components (in consideration of the functionality of the species and habitats). The ecosystem level information and management could be more efficiently addressed by making explicit ecosystem requirements within the scoring of other P2 components (as has been done in P1). For example, 100 level scores in relation to information or management for a particular component would need to show explicit consideration of the wider ecological role of that component. Or consideration of wider ecosystem impacts could be included as a requirement within the definition of 'strategy'.

If the decision is taken that Ecosystem should remain as a stand-alone component, then consideration should then be given to whether the ecosystem expectations can be demonstrated to be met in fewer than 9SIs. In particular the 5 SIs related to ecosystem information (PI2.5.3) are particularly repetitive.

3.2. Changes for the future

The solutions proposed up to now retain the focus on key components and elements and seeks to maintain the intent and the scoring thresholds of the existing standard whist seeking to increase simplicity, clarity and efficiency.

An alternative solution, which would require greater change (and which is therefore not proposed for consideration during the current MSC standard Review) would be to shift the focus away from key components and elements and instead focus on the characteristics of the precautionary management systems. This would seek to make the MSC an audit of the management system and its ability to respond to fishery related environmental impact and deliver positive environmental outcomes.

By moving away from a focus on key elements and components and focusing more on the management system, further questions about the structure of the standard would be likely to arise, including the current separation into 3 Principles.

3.2.1. Are fisheries really divided into 3 Principles?

The 3 MSC Principles are sensible and have proved useful over the years in assessing fisheries. They also often align with the expertise within an audit team, with the P1 assessor being a stock assessment scientist, the P2 assessor being a marine ecologist and the P3 assessor coming from a fisheries management or administration background. In previous standard reviews each Principle has come under review and scrutiny and changes have been made to the content and structure within the Principle. However, typically these Principle reviews have been done in isolation from each other so comparatively little consideration has been given to whether the Principles themselves are still useful and appropriate.

The MSC standard structure is often used as a structure for action-planning within a Fisheries Improvement Programme (FIP). However, the order or chronology of the MSC standard may not be the most useful structure for FIP action planning. For example, a FIP might sensibly begin with consideration of legislative and administrative structures and consideration of roles and responsibilities. It might then follow with a review of data collection systems across all components (both P1 and P2) and a review of management procedures and decision-making processes. Finally, resources could be applied to undertake stock assessments or implement additional management or research on particular components or elements as required. This represents a more holistic approach with a more logical chronology which is not constrained by separation into MSC principles and PIs. A

more hierarchical structure might make more sense from a FIP point of view, whilst still allowing all existing areas within the MSC scoring to be covered in a way which reduces the likelihood of repetition (for example by requiring the fisheries information collection system to be described for every component). A hierarchical structure could begin by:

- consideration of the high level or cross-cutting themes (such as legislation, high level objectives, roles and responsibilities);
- consideration of more process-related themes (such as fisheries information systems, consultation processes, review processes, dispute resolution, control and enforcement or management processes);
- consideration of more component-specific actions and outcomes (such as technical measures, habitat management or stock assessment).

The overall content and scope of the standard as well as the required areas of auditor expertise would remain unchanged but the division between the three discreet principles would cease.

3.2.2. Is it actually all about Management?

In conducting this exercise, the project team attempted to clarify our understanding, based on the existing guideposts, of the definitions of "Measures", "Partial Strategy" and "Strategy". In the table below we repeat the table presented earlier in the report which is derived from consideration of existing definitions.

	Measures	**Partial Strategy**	**Strategy**[18]
Scope	UoA or wider		UoA *and* wider
Objective	Limiting impact / not hindering recovery		Defined Management target
Design	Either designed for component, or incidental		Designed for component
Linkages	Unlinked	Some cohesive links	Strategically linked
Responsiveness	Non-responsive	Response where shown to be ineffective	Fully responsive
Cumulative	Just UoA	UoA and other MSC fisheries	All fisheries
Direct / indirect impacts	Direct only		Direct & Indirect
Monitoring		Some	Full

[18] For ETP there is an additional level of "Comprehensive strategy" however the only distinction with "strategy" is that it is "complete and tested". Although reference is made to monitoring, analysis and responsiveness, these appear to be required at the "strategy level".

The next logical step following on from this exercise is to consider whether further requirements could be included within this definition. For example, we have already suggested that the frequency of review of 'alternative measures' could be usefully added into this framework and that requirements for wider ecosystem considerations could also be included. The current proposal is that all criteria would need to be met in order to meet the defined management level.

We then considered whether this structure could even be further extended to include all remaining areas of the MSC standard and actually become the foundation of a revised overall assessment structure. For example, might the definition of management level also include thresholds for:

- the level of consultation,
- the level of transparency,
- the level of research,
- the effectiveness of dispute resolution,
- the level of control and enforcement,
- the scope of evaluations and review,
- etc.

The three columns of "measures", "partial strategy" and "strategy" could become the basis for revised scoring guideposts. And when considering a species at P1, is it actually just requiring a slightly higher level of management.

So, could an extended and expanded management definition be the basis for scoring all elements of the fishery? Those elements of the catch that scored highest (across all of the requirements of management) might be eligible to be MSC certified (i.e. P1). And those elements of the catch that have areas of management requirements at the lower level would still require a condition.

This more clearly shifts the focus of the MSC onto 'good management', and away from 'outcome status' to become a genuine management audit. This could likely be achieved even with relatively little change to the actual scoring thresholds.

3.3. Comparative Review

A comparative review of the existing and the new proposed structure has been undertaken to illustrate what has been removed or amended from the standard. Because some text has been moved and some SIs combined this does not provide an absolute indication of which text has been deleted or added, but is indicative. This is provided as a separate submission, along with this report.

The image below provides a colour-coded comparison of the existing and the proposed new structure, seeking to illustrate which of the existing SIs are covered by the proposed new SIs.

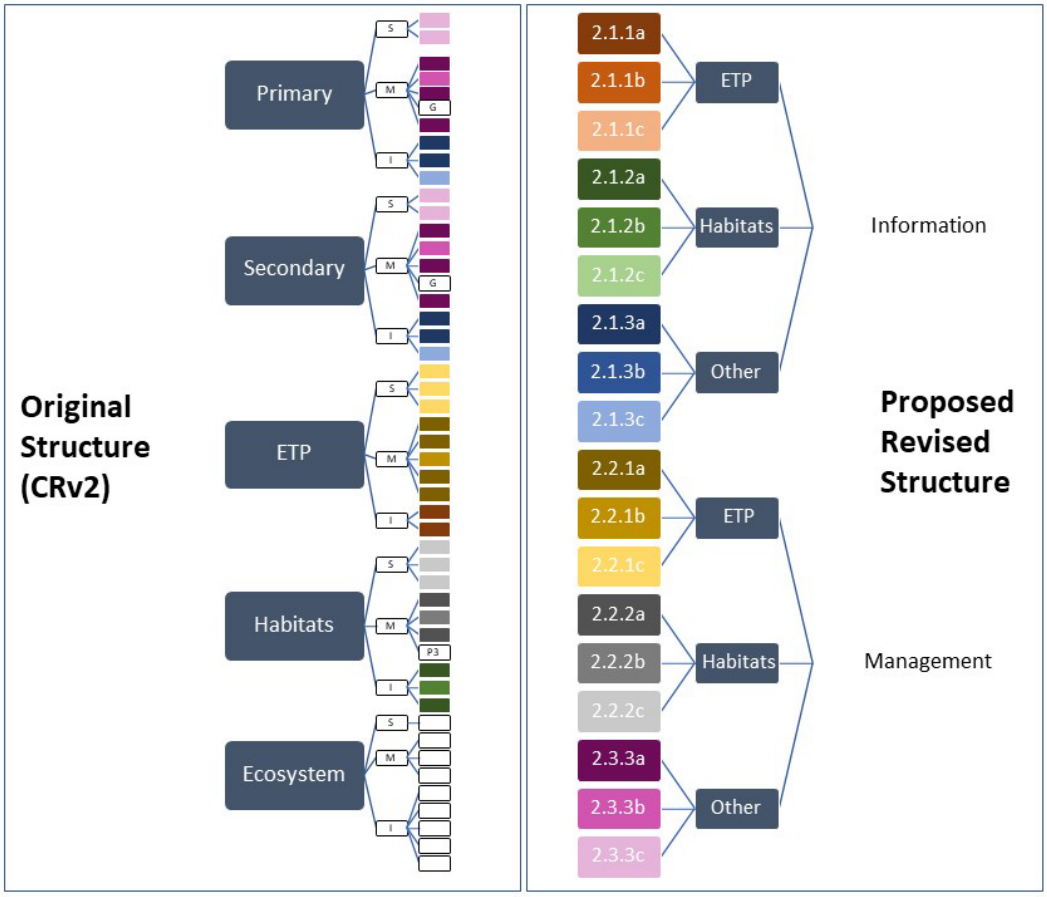

Figure 4: Structure of existing P2 (left) and proposed alternative P2 (right), with colour coding used to illustrate where areas in existing structure are covered in the new structure. S = Status; M = Management; I = Information; G = Proposal to shift SI requirement to Guidance; P3 = Proposal to shift requirement to P3.

3.4. Summary

The proposed changes and the appended revised P2 structure for the MSC assessment tree is not a finished product. It is intended to explore the ideas discussed at the April 2018 efficiency workshop to test the concept. The project team concludes that this concept shows potential to increase efficiency and clarity and is therefore a positive change which should be explored further. The next stage for this concept will be to present it the MSC Technical Advisory Board. If positively received a further of impact testing will be required before the proposal is more widely consulted upon. Even if the proposed revised structure is ultimately rejected the goal of clarity, simplicity and efficiency must continue to prevail in future standard reviews.

Appendix 1: Proposed Revised P2 Assessment Tree Structure

PI	Title	SI	
2.1.1	**ETP Species Information**	A	to assess Impacts
		B	to assess Consequence
		C	to assess Management
2.1.2	**Habitat Information**	A	to assess Distribution and vulnerability
		B	to assess consequence
		C	to assess Management
2.1.3	**Other caught-species Information**	A	to assess Impacts
		B	to assess Status
		C	to assess Management
2.2.1	**ETP Species Management**	A	Strategic level
		B	Evaluation
		C	Outcome
2.2.2	**Habitats Management**	A	Strategic level
		B	Evaluation
		C	Outcome
2.2.3	**Other caught-species Management**	A	Strategic level
		B	Evaluation
		C	Outcome

Comparison of old and new structures

Original Structure				Location within proposed new structure
Component	Performance Indicator (PI)		(SI)	
Primary species	2.1.1	Outcome	a	2.2.3c
			b	2.2.3c
	2.1.2	Management strategy	a	2.2.3a
			b	2.2.3b
			c	2.2.3a
			d	Guidance
			e	2.2.3a (definition of Management level)
	2.1.3	Information	a	2.1.3a&b
			b	2.1.3a&b
			c	2.1.3c
Secondary species	2.2.1	Outcome	a	2.2.3c
			b	2.2.3c
	2.2.2	Management strategy	a	2.2.3a
			b	2.2.3b
			c	2.2.3a
			d	Guidance
			e	2.2.3a (definition of Management level)
	2.2.3	Information	a	2.1.3a&b
			b	2.1.3a&b
			c	2.1.3c
ETP species	2.3.1	Outcome	a / b	2.2.1c
			c	2.2.1c
	2.3.2	Management strategy	a /b	2.2.1a
			c	2.2.1b
			d	2.2.1a
			e	2.2.1a
	2.3.3	Information	a	2.1.1a&b
			b	2.1.1c
Habitats	2.4.1	Outcome	a	2.2.2c
			b	2.2.2c
			c	2.2.2c
	2.4.2	Management strategy	a	2.2.2a
			b	2.2.2b
			c	2.2.2a
			d	P3
	2.4.3	Information	a	2.1.2a
			b	2.1.2b
			c	2.2.2a (definition of management level)
Ecosystem	2.5.1	Outcome	a	2.2.2c (habitats) elsewhere within management SIs
	2.5.2	Management	a	Within definition of management level
			b	Within definition of management level
			c	Within definition of management level
	2.5.3	Information	a	Within SG100 information SIs
			b	Within SG100 information SIs
			c	Within SG100 information SIs
			d	Within SG100 information SIs
			e	Within SG100 information SIs

2.1.1 ETP Species Information

PI 2.1.1		ETP Species Information		
Scoring Issue		SG 60	SG 80	SG 100
A	For estimation of mortality of ETP species			
	Guide post	*Information is adequate to……*		
		… form an expert judgement of the scale of UoA related mortality of ETP (or susceptibility attributes, where RBF is used).	… quantitatively determine the scale of UoA related mortality of ETP (or susceptibility attributes, where RBF is used).	… fully quantify with a high degree of certainty the scale of UoA related mortality and injuries to ETP species.
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	*Data, references, stakeholder comments*		
B	For estimation of consequences to ETP species			
	Guide post	*Information is adequate to……*		
		… form an expert judgement of whether UoA related mortality hinders protection and recovery of ETP species (or productivity attributes, where RBF is used to score outcome status).	… quantitatively determine the consequence of UoA related mortality on the status of ETP species (or productivity attributes, where RBF is used to score outcome status).	… fully quantify with a high degree of certainty the consequences of UoA related mortality on the status and ecosystem function of ETP species.
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	*Data, references, stakeholder comments*		
C	To support ETP management			
	Guide post	*Information is adequate to support……*		
		…. **measures** to manage impacts on ETP species.	….. a **strategy** to manage impacts on ETP species.	…. a **comprehensive strategy** to manage impacts (both direct and indirect) on ETP species.
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
CONDITION NUMBER (if relevant):				

2.1.2 Habitats Information

PI 2.1.2		Habitats Information		
Scoring Issue		SG 60	SG 80	SG 100
A	On habitat types, distribution and vulnerability			
	Guide post	*Information is adequate to …….* …. identify VMEs and broadly identify the nature and distribution of commonly encountered habitats (or habitat types, where RBF is used).	… quantitatively determine nature, distribution and vulnerability of habitats (or habitat types and attributes, where RBF is used).	… fully quantify with a high degree of certainty the nature, distribution, vulnerability and ecosystem function of all habitats.
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
B	For estimation of the consequence to habitat status			
	Guide post	*Information is adequate to ……* … form an expert judgement of the scale of the main impacts of the UoA on VMEs (or consequence attributes, where RBF is used to score outcome status).	…. quantitatively determine the impacts of the UoA on VMEs and commonly encountered habitats (or consequence attributes, where RBF is used to score outcome status).	… fully quantify with a high degree of certainty the impacts of the UoA on all habitats with a high degree of certainty.
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
C	To support habitats management			
	Guide post	*Information is adequate to support…* … **measures** to manage impacts on habitats.	… a **partial strategy** to manage impacts on habitats.	… a **strategy** to manage impacts on habitats.
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
CONDITION NUMBER (if relevant):				

2.1.3 Other caught species Information

PI 2.1.3		Other caught species Information		
Scoring Issue		SG 60	SG 80	SG 100
a	For estimation of mortality of other caught species			
	Guide post	*Information is adequate to …* … form an expert judgement of scale of UoA related mortality of other caught species (or susceptibility attributes, where RBF is used to score outcome status).	…quantitatively determine UoA related mortality of other caught species (or susceptibility attributes, where RBF is used to score outcome status).	… fully quantify with a high degree of certainty UoA related mortality of other caught species.
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
b	For assessment the consequence to other caught species			
	Guide post	*Information is adequate to …* … form an expert judgement of the consequence of UoA related mortality on recovery or rebuilding of other caught species (or productivity attributes, where RBF is used to score outcome status).	… quantitatively determine consequence of UoA related mortality on the status other caught species (or productivity attributes, where RBF is used to score outcome status).	… fully quantify with a high degree of certainty the consequence of UoA related mortality on status and ecosystem function of other caught species.
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
c	To support management of other caught species			
	Guide post	*Information is adequate to support…* … **measures** to manage other caught species.	… a **partial strategy** to manage other caught species.	… a **strategy** to manage other caught species**.**
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
CONDITION NUMBER (if relevant):				

2.2.1 ETP Species Management

PI 2.2.1		ETP Species Management		
Scoring Issue		SG 60	SG 80	SG 100
a	**Management strategy in place**			
	Guide post	*In order to manage impacts on ETP species …* **… measures** are implemented.	… a **strategy** is implemented.	… a comprehensive **strategy** is implemented (designed to exceed national or international requirements).
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
b	**Management strategy evaluation**			
	Guide post	*The effectiveness of management has been determined by ….* … expert judgement (e.g. general experience, theory or comparison with similar UoAs/species).	… a review (using some information directly about the UoA and/or species involved)	… testing and quantitative evaluation (using information directly about the UoA and/or species involved)
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
c	**Management strategy outcome (status)**			
	Guide post	Direct effects of the UoA are unlikely to hinder recovery and rebuilding of ETP species. Or RBF Score	Direct and indirect effects of the UoA and other MSC UoAs are unlikely to hinder recovery and rebuilding of ETP species. Or RBF Score	The UoA and other MSC UoAs are unlikely to have significant detrimental effects (either direct or indirect) on ETP Species. Or RBF Score
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
CONDITION NUMBER (if relevant):				

2.2.2 Habitats Management

PI 2.2.2		Habitats Management		
Scoring Issue		SG 60	SG 80	SG 100
a	**Management strategy in place**			
	Guide post	*In order to manage impacts on habitats …* **… measures** are implemented, if necessary.	… a **partial strategy** is implemented, if necessary.	… a **strategy** is implemented.
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
b	**Management strategy evaluation**			
	Guide post	*The effectiveness of management has been determined by ….* … expert judgement (e.g. general experience, theory or comparison with similar UoAs/species).	… a review (using some information directly about the UoA and/or species involved).	… testing and quantitative evaluation (using information directly about the UoA and/or species involved).
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
c	**Management strategy outcome (status)**			
	Guide post	The UoA is unlikely to hinder the recovery and rebuilding of VMEs. Or RBF Score	The UoA and other MSC UoAs are unlikely to reduce the structure and function VMEs and commonly encountered and vulnerable habitats to the point where there would be serious or irreversible harm. Or RBF Score	The UoA and other MSC UoAs are likely to have no significant detrimental effects on habitat structure and function or ecosystem role. Or RBF Score
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
CONDITION NUMBER (if relevant):				

2.2.3 Management of other caught species

PI 2.2.3		Other caught species management		
Scoring Issue		SG 60	SG 80	SG 100
a	**Management strategy in place**			
	Guide post	In order to manage impacts on other species … **… measures** are implemented, if necessary.	… a **partial strategy** is implemented, if necessary.	… a **strategy** is implemented.
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
b	**Management strategy evaluation**			
	Guide post	*The effectiveness of management has been determined by ….* … expert judgement (e.g. general experience, theory or comparison with similar UoAs/species).	… a review (using some information directly about the UoA and/or species involved).	… testing and quantitative evaluation (using information directly about the UoA and/or species involved).
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
c	**Management strategy outcome (status)**			
	Guide post	If below PRI (or biologically based limits where PRI is undefined) the UoA is unlikely to hinder recovery and rebuilding of other caught-species. Or RBF score	If around or below the PRI (or biologically based limits where PRI is undefined) the UoA and other MSC UoA are unlikely to hinder recovery and rebuilding of other caught-species Or RBF score	Other caught species are likely to be fluctuating around or above MSY or proxy. Or RBF score
	Met?	(Y/N)	(Y/N)	(Y/N)
	Justification			
	Audit Trail	Data, references, stakeholder comments		
CONDITION NUMBER (if relevant):				