An Assessment of Japanese Whale Research Programs Under Special Permit in the Antarctic (JARPA, JARPA II) as Programs for Purposes of Scientific Research in the Context of Conservation and Management of Whales

Marc Mangel, PhD, FAAAS, FRSE

University of California

Santa Cruz

About the Author

Marc Mangel, PhD, is Distinguished Professor of Applied Mathematics and Statistics and Jack Baskin Endowed Chair in Technology and Information Management at the University of California Santa Cruz (UCSC). He also directs the Center for Stock Assessment Research, a cooperative program between the Fisheries Ecology Division, Southwest Fisheries Science Center, NOAA Fisheries and UCSC to train students and post-doctoral scholars in the methods of quantitative population biology required for sustainable fisheries. His honors include Fellowship in the American Association for the Advancement of Science and the Royal Society of Edinburgh. Further biographic details are given in Appendix A.

TABLE OF CONTENTS

1. Executive summary 337

2. Introduction 339

3. An overview of whaling in the antarctic 340

4. Characteristics of a program for purposes of scientific research .. 349

5. Description and assessment of JARPA and JARPA II
as programs for purposes of scientific research in the conext of
conservation and management of whales 360

6. Conclusion 376

7. Literature cited... 379

8. Appendices 385

1. EXECUTIVE SUMMARY

1.1. In this Expert Opinion, I provide an assessment of the Japanese Whale Research Program under Special Permit in the Antarctic (JARPA) and JARPA II (the second phase) as programs for purposes of scientific research in the context of conservation and management of whales based on generally accepted scientific practice and criteria developed by the International Whaling Commission (IWC) for Special Permit research.

1.2. I begin with an overview of whaling in the Southern Ocean, as it pertains to the assessment of JARPA and JARPA II as programs for purposes of scientific research in the context of conservation and management of whales. I explain how the Revised Management Procedure (RMP) of the IWC is an advance in management ideas that allows effective conservation and management of whales without detailed biological knowledge and without use of lethally obtained data.

1.3. Consistent with generally accepted scientific practice and with criteria identified by the Scientific Committee of the IWC, my opinion is that the essential characteristics of a program for the purposes of scientific research in the context of conservation and management of whales are that the program:

a) has defined and achievable objectives that aim to contribute knowledge that is important to the conservation and management of whale stocks;

b) employs appropriate methods that are likely to achieve the stated objectives, including:

   (i) lethal methods only where the objectives of the research cannot be achieved by any other means (for example, by the analysis of existing data and/or the use of non-lethal research techniques);

   (ii) setting sample sizes using accepted statistical methodology; and

   (iii) linking mathematical models to data consistently;

c) includes periodic review of research proposals and results and adjustment in response to such review; and

d) is designed to avoid adverse effects on the stocks being studied.

1.4. I then assess JARPA and JARPA II against those criteria and conclude that they meet none of them.

1.5. First, JARPA II does not - and JARPA did not - have defined and achievable objectives that aim to contribute knowledge that is important to the conservation and management of whales. JARPA II has - and JARPA had - broad and vague objectives that conflate exploration and exploitation. Their stated objectives could be used to justify almost any activity that Japan wished to pursue. Their contribution to management remains undemonstrated after 24 years and the potential of JARPA II to bring new knowledge about the conservation and management of whales is very low, if it indeed exists at all.

1.6. Second, JARPA II does not - and JARPA did not - employ appropriate methods likely to achieve its stated objectives. Although a variety of empirical methods are in principle employed in JARPA II, a majority of effort is devoted to lethal take despite the existence of problems with the data generated by that lethal take and despite the existence of other, non-lethal, methods that can provide nearly all of the same information. The reasoning that underlies the setting of sample sizes (the number of animals killed) and the distribution of sampling effort is vague, unclear, and at times simply wrong. The links between the proposed models of the ecosystem and the field work, particularly lethal take, are weak and unclear.

1.7. Third, most of the work done in association with JARPA and JARPA II is published outside of standard peer-reviewed literature. Only about 15% of the published papers are peer-reviewed and potentially relevant to the stated objectives. Workers in JARPA and JARPA II have not demonstrated an ability to respond to criticism or to admit being wrong.

1.8. Fourth, there is no record of any attention being directed to avoiding unintended adverse consequences in the design of JARPA or JARPA II; indeed they proceed on the assumption that the take will have no effect on the stock.

1.9. My conclusion is that JARPA II is - and JARPA was - an activity for the collection of data in the Southern Ocean. However, both have failed at turning data into knowledge or in improving the conservation and management of whales. JARPA II is not a program for purposes of scientific research in the context of conservation and management of whales.

2. INTRODUCTION

2.1. I have been asked by the Government of Australia to prepare an independent report on the Second Phase of the Japanese Whale Research Program under Special Permit in the Antarctic (JARPA II) and related matters. The full terms of reference provided to me are found in Appendix B. Briefly, they are:

- *To identify and outline the essential characteristics of a program undertaken for purposes of scientific research; and*
- *To provide a critical analysis of the objectives, methodologies and other features of JARPA II and, in doing so, assess whether JARPA II has the essential characteristics of a program undertaken for purposes of scientific research.*

2.2. I was given background material briefly described in **Appendix C**.

2.3. In order to meet the terms of reference, it is essential to understand the characteristics of a program for purposes of scientific research in general and in the specific context of conservation and management of whales. It is also essential to understand the nature of JARPA II (and its predecessor JARPA), so that they can be assessed as to whether they may properly be characterized as programs for purposes of scientific research in the context of conservation and management of whales.

2.4. In this paper, I

a) give a brief overview of whaling in the Antarctic, emphasizing the key points that are relevant for the subsequent analysis;

b) identify the essential characteristics of a program for purposes of scientific research in general and in the specific context of conservation and management of whales;

c) provide an overview of the relevant aspects of JARPA and JARPA II, and assess them against the essential characteristics of a program for purposes of scientific research in the context of conservation and management of whales; and

d) conclude with a summary of that assessment.

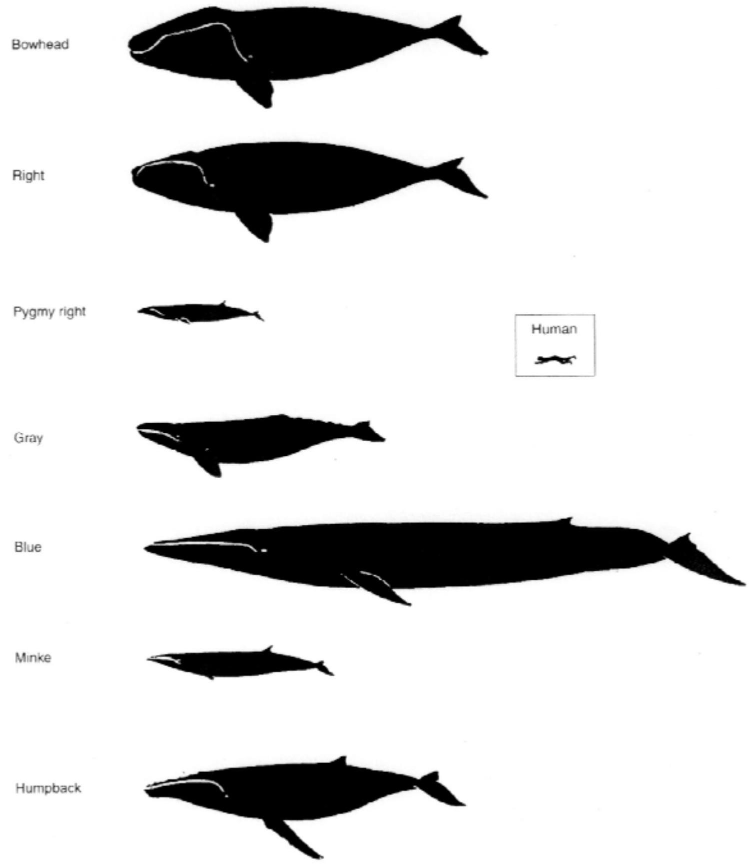
3. AN OVERVIEW OF WHALING IN THE ANTARCTIC

3.1. Modern commercial whaling began early in the 20th century, using land-based stations (Mackintosh 1965). The first Antarctic whaling station was established at South Georgia in 1904. Whaling at the South Shetland and Orkney Islands almost exclusively used factory ships, which were tankers fitted with a factory plant and moored in a harbor to function as a floating land station. Land-based whaling ran from about 1904 to 1928, after which the great era of pelagic (at sea) whaling followed. By 1930/31, there were 41 pelagic factories with over 200 catching vessels working in the Antarctic.

3.2. However, during the 1930/31 Antarctic whaling season, more whale oil was produced than the world market could absorb. Because of this, the whaling companies agreed to limit their output and devised a plan to regulate catches by the amount of oil produced. Since the species of major commercial take in the early 20th century were the blue, fin, sei, and humpback whales (Mackintosh 1965), an effort was made to put them into a common currency. One blue whale was considered the same as 2 fin, 2.5 humpback, or 6 sei whales; giving rise the notion of the Blue Whale Unit (BWU) (Gambell 1999, Gillespie 2005).

3.3. In the figure below (from Bannister 2002), I show the relative sizes of some of the whales.

3.4. The blue, fin, humpback and sei were called the great whales. At the time that the BWU was conceived, minke whales were not considered relevant to commercial whaling because of their small size. Tønnessen and Johnsen (1982) noted that, had minke whales been considered, one BWU would have been at least 30 and possibly up to 60 minke whales.



The International Whaling Commission

3.5. The inter-governmental International Whaling Commission (IWC) (Gambell 1999, Donovan 2002) is charged, among other things, with regulating whaling in the Southern Ocean. The IWC was established in 1946 through the International Convention for the Regulation of Whaling (ICRW). The ICRW consists of two parts: the convention itself and a schedule of regulations intended to govern whaling operations. Contracting Parties to the ICRW subscribe to:

a) safeguarding for future generations the great natural resources represented by whale stocks;

b) protecting all species of whales from further over-fishing;

c) seeking the optimum level of whale stocks;

d) providing an interval for recovery to certain species of whales now depleted in numbers; and

e) establishing a system of international regulation for the whale fisheries to ensure proper and effective conservation and development of whale stocks (Gillespie pg 396-397).

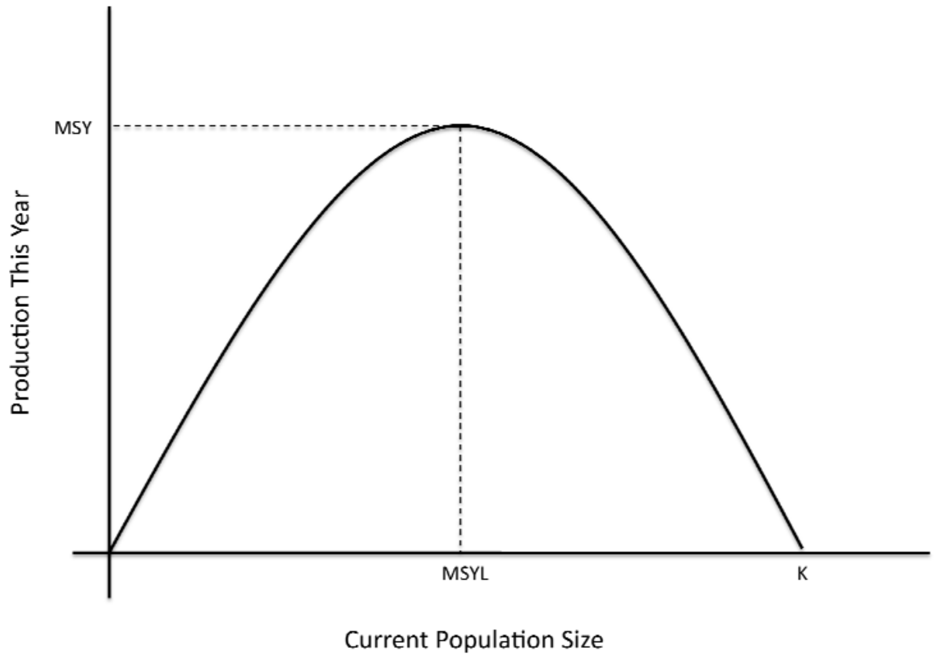
3.6. The IWC has a Scientific Committee that meets annually, usually for two weeks before the annual Commission meeting, and that often also holds *ad hoc* meetings between sessions.

3.7. From its inception until about 1972, the IWC regulated whaling using the BWU. The overall catch limit was initially set to 16,000 BWUs, with no reference to specific species except that some species (e.g. right whales, humpback whales) were designated as protected. This was essentially an open access fishery (as defined by Clark 2006), in which nations raced to catch as many whales as possible before the quota was reached, leading to waste during processing, an uneconomical increase in the number of catcher boats, and poor conservation of the whales (Donovan 2002). Furthermore, the quotas were often exceeded. By 1952 it was recognized that there were problems with this management procedure, and in 1963 a small group of eminent scientists appointed by the IWC recommended elimination of the BWU as a method of setting catch limits (Clapham and Baker 2002). By 1971/72 the catch limit had been reduced to 2,300 BWUs and both blue and humpback whales had been protected from commercial whaling.

Fundamentals of the Dynamics of Populations

3.8. I now briefly describe how the dynamics of populations are characterized and some of the terminology used in the management and conservation of whales.

3.9. In population dynamics, a key focus is the change in population size from one year to the next. This is called net production and is measured most simply as the number of individuals in one year minus the number of individuals in the previous year. Such net production usually depends upon current size of the population and is typically a peaked function of current population size as shown in the figure below:



3.10. When there are no individuals, in the absence of immigration, net production is 0 since without any individuals no new individuals can be produced.

3.11. When there are many individuals (indicated by *K* on the x-axis in the figure) net production is also zero because competition for food causes a balance between births and deaths. (The rates of birth and death, the latter commonly called the rate of mortality, are called the demographic (or biological) parameters of the population.) When such balance between births and deaths is achieved, the population is in a state known as carrying capacity. In the absence of environmental fluctuations, this is the size at which the population would stabilize if it were to remain unexploited.

3.12. The population size that maximizes net production is called the Maximum Sustainable Yield Level (MSYL) and the level of production associated with that population size is called Maximum Sustainable Yield (MSY). Any catch that is higher than the MSY indicated by the peak of the curve in Para 3.8 is not sustainable, since more is being taken from the population than is being produced by it. The MSY Rate (MSYR) is the ratio of MSY to MSYL. For many years, and continuing to this day in some cases, managing fisheries for MSY was a standard approach.

3.13. Ricker (1975) defined MSY as "[t]he largest average catch or yield that can continuously be taken from a stock under existing environmental conditions." Ricker's definition hinges on three key words: average, continuously, and existing. That is, in nature there is not a single curve as shown in Para 3.8 but a family of such curves, depending upon environmental (both physical and biological) conditions. For example, the changing biomass of krill as water temperature changes will affect the carrying capacity for whales (Wiedenmann et al 2008). As environmental conditions vary, the shape of the curve, location of carrying capacity (*K*) and the value of MSY may also vary. Moreover, if one does not know the curve precisely and one does not know the current population size precisely, then one never knows that the catch is sustainable even assuming that environmental conditions are constant. Consequently, uncertainty and environmental variation make MSY a fragile concept for management purposes.

3.14. Larkin (1977) argued that MSY should be put to rest because, among other things, it led to yields that were too high and unsustainable. He wrote "[w]hatever lies ahead in the development of new concepts for harvesting the resources of the world's fresh waters and oceans, it is certain that the concept of maximum sustained yield will alone not be sufficient" (Larkin 1977, 10). That is, MSY should be considered a constraint rather than a target since harvests greater than MSY are not sustainable (Mangel et al 2002).

The New Management Procedure

3.15. After 1972, the IWC abandoned the BWU and in 1974 adopted a realignment of its management procedures through the development of a New Management Procedure (NMP). The NMP was designed to calculate catch limits for whale populations using the fundamental principles of population dynamics as described in Paras 3.8-3.14. The goals of the NMP were to bring each of the whale stocks to the particular population level at which MSY could occur and to protect stocks whose population sizes were estimated to be below a fixed fraction of their pre-industrial exploitation level (Gambell 1999, Donovan 2002).

3.16. The NMP aimed to separate stocks into the three categories, based on the extent to which the size of the stock varied from MSYL (Gillespie 2005):

a) initial management stocks (those considered to be above the size generating MSY and which could thus be harvested down to that level);

b) sustained management stocks (which were close to the size generating MSY and would be maintained there); and

c) protection stocks (those that would not be harvested).

3.17. However, over time it became clear that the NMP had serious problems (Cooke, 1995; de la Mare 1986abc, Holt 2004). The NMP was based on MSY, although at the time it was proposed, the data required to calculate MSY were lacking. Thus, two *ad hoc* rules were added. First, stocks that had been subject to stable catches over considerable periods of time would continue to be harvested as long as there was no evidence of a decline. Second, for stocks that had not been subject to serious previous exploitation, catches would be limited to 5% of the estimated size of the stock. This rule was precautionary, in the sense that before harvest began, population estimates had to be obtained. However, the NMP did not deal with the question of how to incorporate the uncertainty in the estimates of population size. Indeed, one of the failings of the NMP was that it did not stipulate how existing data were to be used to assess the state of the stock and it could not handle uncertainty regarding the status of the stock in a robust manner (Cooke 1995, pg 652).

3.18. Cooke (1995, pg 648) noted

> The main difficulty in operating the NMP was that there were insufficient data for its implementation. For most stocks there was no reliable estimate of population size, let alone an estimate of the MSY or the relation between the current population and the MSY level. Furthermore there was no particular incentive to collect data. Even if relatively good data had been available, there would still have been considerable uncertainty about the state of whale stocks with respect to the NMP criteria, but there were no guidelines as to how to cope with these uncertainties. Finally, the 'behaviour' of the procedure was unknown. By this is meant the expected long-term consequences of applying the procedure to whale stocks.

Indeed, it was still possible for whale stocks to be depleted even if the NMP were followed precisely and the population dynamics of the whales matched those assumed in the NMP because of the uncertainty associated with estimates of population size (Cooke 1995).

The 1982 Moratorium on Commercial Whaling

3.19. The IWC adopted the moratorium on commercial whaling in 1982, setting catch limits for all stocks at zero with effect from the 1986 coastal and the 1985/86 pelagic seasons. The commercial whaling moratorium remains in force today.

3.20. One of the objectives of the decision to institute the moratorium was to provide time for the IWC to establish its best estimate of population sizes together with a suitable procedure to facilitate sustainable catch limits. A moratorium on commercial whaling would also enable the IWC to develop mechanisms by which the whaling industry could be effectively regulated to avoid the problems of the past.

The Revised Management Procedure (RMP)

3.21. Because of the serious problems with the NMP, the IWC spent about a decade developing a Revised Management Procedure (RMP). The specific goals of the RMP are:

a) to achieve stable catch limits, thus allowing the orderly development and regulation of the whaling industry;

b) to manage acceptable risk and to ensure that a stock is not depleted to the point where the risk of extinction is not negligible; and

c) to ensure the highest possible continuing yield from each whale stock.

3.22. In order to achieve the goals of the RMP, the IWC agreed that (IWC 1994, 1999):

a) commercial whaling would be permitted only for populations in areas and seasons for which catch limits (calculated by its Scientific Committee and approved by the Commission) were in force;

b) catches would reach a maximum level when a stock was at 72% of its unexploited level; and

c) there would be no whaling on stocks that were below 54% of their unexploited level.

3.23. The aim of the RMP is not to attempt to calculate MSY or any other optimum level. Rather, it is intended to effectively manage whaling while dealing with the inherent uncertainty in the Southern Ocean ecosystem. Holt (2004, pg xii-xiii, italics added) described the RMP in this manner:

> Although the RMP uses a population model for the estimation of stock status and the calculation of catch limits, the model itself is hugely simplified. *It does not attempt to emulate the dynamics of any real whale population, and, in fact, does not even explicitly include demographic parameters such as natural mortality rate*. Rather, the simple model is part of a freely invented algorithm that has been shown, by simulations, to meet the targets efficiently and to be robust to errors and such things as environmental changes (Holt 2004, pg xii-xiii, emphasis added).

3.24. In much the same way as a good card player will compute the odds that an opponent has a certain card, under the RMP statistical methods are used to produce a probability distribution for the catch limit and the current population size, which is measured as a fraction of the unexploited level. Catch limits are computed using a Catch Limit Algorithm (CLA), which sets the catch limit to be 0 if the population abundance is estimated to be less than 54% of its unexploited level. If population abundance is estimated to be more than 54% of unexploited level, then the catch limit is set at a specified fraction of the population above the unexploited level.

3.25. The data used in the CLA comprise only:

a) total catch statistics based on previous whaling (past data); and

b) data obtained through sighting surveys in which ships follow a prescribed track line and count the number of whales that are seen (current and future data).

3.26. The RMP thus eliminates the use of data obtained from whaling-dependent or other lethal-source data, which are often unreliable for purposes of management because they represent non-random samples of the population. Consistent with this, in 1995 the IWC adopted a Resolution (1995-9) that stated, among other things, “that scientific research intended to assist the comprehensive assessment of whale stocks and the implementation of the Revised Management Procedure *shall be undertaken by non-lethal means*” (emphasis added).

3.27. The development of modern computational tools, particularly the capacity to undertake extensive computer simulation, allowed thorough testing of the RMP (Kirkwood 1992, Cooke 1995). That is, the RMP was tested using sets of pseudo-data that had been generated by other, more complicated population models. The point of this testing was to ask the question: “how effective is the RMP in setting catch limits that maintain or restore populations to acceptable levels when various demographic parameters are unknown, or when the structure of the actual population dynamics differs from those assumed in the RMP?”

3.28. The tests allowed assessments of the performance of the RMP with incorrect assumptions about the dynamics of the stock, varying initial abundance, bias in sighting surveys, different relationships between true abundance and catch per unit effort (a common proxy for abundance), uncertain or inaccurate catch histories, and/or rare episodic events (e.g. epidemics). The tests showed that the RMP was robust to these variations, maintaining catch and preventing the depletion of the population (Cooke 1995). Most importantly, the tests allowed the IWC to conclude that the RMP functioned effectively without making specific assumptions about the population dynamics of whales and taking into account possible errors in historic catch record.

3.29. In contrast, I know of no peer-reviewed published paper that demonstrates fundamental flaws with the RMP that can only be corrected through field-based programs that involve lethal take.

3.30. In conclusion, the RMP is an "advance in management ideas" (Holt 2004, pg xiii) and consistent with other scientific work on the most effective level of complexity for models used in management of living marine resources (Ludwig and Walters 1985, Hilborn and Mangel 1997). As a member of the Committee of Scientific Advisors of the U.S. Marine Mammal Commission 1989-1996, I observed but did not participate in the development of the RMP. Returning to it now after a 15 year absence, I am able to assess it with a fresh viewpoint and concur with Holt that it is indeed a substantial advance in management.

3.31. In summary

- The Southern Ocean ecosystem is characterized by uncertainty in many dimensions including the dynamics of populations.
- The commercial whaling moratorium, effective since 1986, has allowed the IWC to develop and test the Revised Management Procedure (RMP) as an effective tool for the future management of whaling.
- The RMP
    - uses an intentionally simple model of population dynamics;
    - is designed so that lethally obtained data are not required;
    - is designed to encourage the collection of sighting information; and
    - has been rigorously tested and found to be robust to variations from its assumptions.

4. CHARACTERISTICS OF A PROGRAM FOR PURPOSES OF SCIENTIFIC RESEARCH

Science as a Process

4.1. The goal of science is to understand the natural world by providing a framework to account for observations already taken and to make predictions of new observations. This goal is achieved by putting new knowledge in the context of existing knowledge, recognizing that even when there is progress the conclusions are transient (that is, subject to ongoing testing and revision) but the methods are not. It is ongoing testing that is the basis for the self-correcting nature of science. Without that self-correction, one cannot claim to be doing science.

4.2. Modern science is complex, and this has led to the 'cult of the expert' (Jenkins 2004). However, much of the complexity can be understood without extensive technical training if one focuses on the characteristics of science as a process for converting data into knowledge. Jenkins (pg 6) wrote "[t]he essence of science is not some nuggets of information about the natural world but rather an ongoing process for gradually learning how the world works, with occasional breakthroughs in the form of major discoveries. At any given time, the understanding of a phenomenon is likely to be incomplete, with conflicting explanations and evidence. Scientists have learned to tolerate such uncertainty and even relish the challenges it offers."

4.3. There are essentially two types of science: (i) textbook science (which most people learn in school) and (ii) science as practiced by scientists, or 'frontier science' (Pickett et al 2007).

4.4. Textbook science is typically identified with the notion of 'the scientific method', which involves:

a) devising alternative hypotheses;
b) devising an experiment (or several of them) with alternative possible outcomes; each of which will, as nearly as possible, exclude one or more of the hypotheses;
c) carrying out the experiment so as to get as clear a result as possible; and
d) recycling the procedure, making sub-hypotheses or sequential hypotheses to refine the possibilities that remain.

In textbook science, we repeatedly challenge a hypothesis with experiments, and if the hypothesis stands up to repeated experiments, it is treated as if it were true.

4.5. Textbook science is a simple, linear process; it is also a myth (Grinnell 2009, l. 70). Science as practiced is more complicated than this and the path to discovery is more convoluted. In the ecological sciences in particular, it is often impossible to conduct experiments, but observation can substitute for experiment (Mangel 2010). In consequence, scientists proceed by assembling many different strands of evidence, which, if collected properly, can be woven into a strong and intellectually sound fabric of conclusions.

4.6. Whether it is textbook or frontier, science does not consist of simply accumulating data. Indeed, we now often face the problem of data 'poisoning' by having too much data and too little understanding. Valiela (2001, pg 11) noted "[d]escription is not tantamount to understanding: descriptive data can not by themselves furnish an explanation of the mechanisms behind the observations, nor can they easily identify the processes that brought about the situation described. Complicated descriptions can become goals in themselves and may delude us into thinking progress has been made". Gopnik (2009, pg 71) noted "[a]ll seeing is impregnated with thinking. If science were simply a bucket into which descriptions fell, it would be a heap of facts. It is in the jump beyond, to a general rule, a theory, even a vision, that science advances".

4.7. Science as practiced by scientists invariably involves weaving many strands of data together to produce new knowledge. The way that this is done depends upon the problem that is being studied, especially in complicated ecological situations where experiments are difficult to impossible. Simply put, the essence of science is to extract knowledge from data and, if one does not know in advance how the data will be analyzed to extract such knowledge, one is not ready to collect the data.

4.8. In accord with generally accepted principles of scientific practice (Valiela 2001, Jenkins 2004, Pickett et al 2007) a program for purposes of scientific research:

a) Has an over-arching conceptual framework that leads to a set of focused questions (hypotheses);

b) Employs the correct set of empirical tools to answer the questions including setting sample sizes with sound statistical reasoning, and linking mathematical models and data appropriately;

c) Has proper assessment through the community of scientists; and

d) Is designed to avoid unintended negative ecological consequences.

An Over-Arching Conceptual Framework Leading to a Set of Focused Questions (Hypotheses)

4.9. A program for purposes of scientific research requires an over-arching conceptual framework. Without it, one simply does 'exploratory analyses' hoping that something interesting will arise from random activity. This rarely works; the Nobel-prize winning immunologist Peter Medawar once wrote "[n]o new principle has emerged from a heap of facts". When people speak of "Newton's Theory of Gravity" or "The Theory of Relativity" or "Darwin's Theory of Evolution by Natural Selection" they mean such overarching conceptual frameworks.

4.10. The conceptual framework provides a clearly specified context and purpose and sets the ground for clearly defined and achievable objectives, but it does not itself lead to a program of work. Rather, it inspires and frames the investigation of particular questions and hypotheses.

4.11. To be testable, questions and hypotheses must be operationally defined – that is, it must be possible to answer the question using existing empirical or theoretical methods or there must be excellent prospect that new methods can be developed to answer the question. For example, Valiela (2001, pg 6) noted "[i]t is a waste of time, of course, to worry about the density of angels on any surface, let alone the head of a pin, unless we have a working seraphometer available".

4.12. Any idea that cannot be operationally defined cannot be studied by empirical science. Similarly, objectives that cannot be tested are not scientific and thus not achievable. In many ecological settings, unlike textbook science, hypotheses may not be mutually exclusive in that an observation clearly excludes hypothesis "A" but not "B". However, even in this case if the hypotheses are operationally defined it is possible to test them and assess the relative strength of the hypotheses provided by the data (Hilborn and Mangel 1997, Wolf and Mangel 2008).

4.13. In the ecological sciences, it is often impossible to conduct experiments. For example, it is not possible to undertake experimental manipulation when attempting to understand the dynamics of populations of blue whales. That is, there is no possibility to replicate an experiment, since there are so few individuals, those individuals may actually constitute a single population, and the time scale of their population dynamics is very long. Nevertheless, we are not prevented from asking questions about blue whales and observation provides a viable means by which to attempt to answer those questions (see for example, Branch et al 2004 on blue whales, Mangel 2010 on Steller sea lions).

The Correct Set of Tools

4.14. Once a set of questions has been established, a program for purposes of scientific research should focus on the important step of identifying the best tools that will answer those questions as clearly and unambiguously as possible. These tools should be selected following an evaluation of their effectiveness in achieving the stated objectives.

Setting Sample Sizes

4.15. Setting the size of a sample of data to be taken in order to estimate an unknown parameter so as to test a hypothesis depends on:

a) how accurately the parameter needs to be known (how close the average value of the estimate is to the unknown parameter);

b) how precisely it needs to be known (how much variation surrounds the estimate of the average value); and

c) what kinds of statistical assessments will be done with the data.

Formal statistical methodology provides procedures by which the sample size required to obtain a specified confidence that we have in a particular conclusion can be determined.

Uses Models Appropriately

4.16. Models have become a cornerstone for extracting knowledge from data. A model is a stylized description used in analyzing or explaining a phenomenon. A model is not a hypothesis in itself. Models are rather tools used in the evaluation of hypotheses. Models serve a number of purposes, one of which is to assist in determining what needs to be measured and how accurate and precise the measurement needs to be. Models that are purported to be linked to field programs must be consistently and appropriately connected to the data from the field program.

Proper Assessment through the Community of Scientists

4.17. Scientists form communities and networks that link to the past and provide connectedness in the present. Grinnell (2009, l.158) noted "[e]ach researcher or group of researchers initiates work in the context of prevailing experiences and beliefs – the starting point and justification for further action". That is, individuals in a program for scientific purposes collaborate in a self-correcting community. Even the greatest geniuses of science (Newton, Darwin, Einstein) had networks and communities and made numerous corrections in their research programs.

4.18. Scientific debate and disagreement is good if it leads to questions that can be resolved by reliable research. Views that are not debatable because they are based on immutable assertions are not scientific since self-correction is not possible. A community with a free exchange of ideas allows scientists to identify occasions when they may be wrong in their research and affords them the opportunity to change their minds. Indeed, delight in the unexpected is the lifeblood of science: "[a]lmost alone in belief systems, science welcomes the disturbingly new" (Raymo 1991, pg 179). Grinnell (2009, l. 385) noted "[t]hey [scientists] are open to the possibility of being wrong". Responding to critical comments and changing research paths is an essential part of the practice of science. An individual who is not open to the possibility of being wrong cannot be a scientist. Furthermore, the scientific community is obliged to expose assumptions, whether they arise from within science or from society, and to explore the implications of those assumptions as they affect the practice of science (Pickett et al 2007).

4.19. That is, scientists belong to a community of independent thinkers cooperating in a relatively free spirit so that a series of independent initiatives becomes organized into joint achievement by "mutually adjusting themselves at every successive stage to the situation created by all the others are who acting likewise" (Polanyi 1969, pg 51).

4.20. Individual scientists sit at the nexus of the world to be studied (in which discovery is the objective) and the research community (in which credibility is the metric). The individual scientist investigates the world and when he or she believes that a discovery is made, the process of conversion from discovery to credibility begins (Grinnell 2009, l. 83). The community of scientists is responsible for the proper assessment and quality control of scientific ideas, in which discovery becomes credibility, through the process of peer review.

4.21. Peer review is a key component for the assessment of the value of ideas (Resnik 2011) and is essential because when the value of an idea is undermined it must be rejected. Peer review both provides quality control on the level of standards of scholarship and methodology for the scientific community and it helps authors improve their research proposals and resulting manuscripts. Peer-review also leads to the generation and establishment of scientific opinion (Polanyi 1969), which is held not by a single individual, but by a collection of individual scientists each of whom endorses the opinions of others. Of course, scientific opinion can be wrong, but reliable science responds to valid criticism, which is how science advances.

4.22. Peer-review requires a multi-dimensional approach for both proposals for research and manuscripts describing the results of research (Polyani 1969). For problems in applied ecology, at the minimum peer-review assesses:

a) plausibility of an idea;

b) scientific value of an idea, consisting of accuracy, intrinsic interest, and importance;

c) originality of an idea, (which is often assessed by the degree of surprise brought about by the idea); and

d) applicability of an idea, assessed by how the work can inform the motivating applied question.

Plausibility and scientific value encourage conformity whereas originality encourages creative thinking and dissent. Applicability ensures that the ideas and the work contribute to solving the motivating applied problem.

4.23. The criteria in Para 4.22 can be melded into questions typically asked by referees assessing proposals for research (Grinnell 2009, l. 332):

a) Is there is a question to be answered?

b) Can the research group answer it?

c) Will getting the answer will be worth the effort?

Unless the answers to all of these questions are “yes”, work should not begin.

4.24. For publications after work has already been done, the questions typically asked by referees are (Grinnell 2009, l. 715):

a) Are the techniques appropriate?

b) Could any scientist potentially have done the work?

c) Are the results interpreted in an appropriate fashion?

d) Are the studies reasonable in light of ideas previously accepted by the community?

Unless the answers to all of these questions are “yes”, the article should not be published.

4.25. Articles that are not peer-reviewed are considered to be ‘grey literature’ and are given less weight than those that have survived the peer-review process. As retractions in high profile journals show, peer review is not perfect, but it is nevertheless an essential characteristic of the practice of science.

4.26. In summary, it is essential to a program for purposes of scientific research that there be peer review from the outset of the research program (since a program should not begin until it has been assessed as feasible through a matching of methods and objectives); that there be peer review throughout the operation of the program (since throughout its duration a program should respond to deviations from objectives by adjusting methods or even abandoning the program in the face of inadequate progress); and that the program end with publication of results in peer-reviewed literature (since it is through peer-reviewed publication that claims of discovery are given scientific credibility).

Is Designed to Avoid Unintended Negative Consequences

4.27. The history of human interaction with the natural world is replete with examples in which human interventions have led to unexpected and surprising consequences. Some of the best examples include those involving the resistance of bacteria to antibiotics and of insects and weeds to pesticides or herbicides respectively.

4.28. Scientific research may have unintended consequences that increase the chance that the population being studied will decline or possibly become extinct. For example, Harrison et al (1991) concluded that their very study of a population of butterflies in California may have lead to its extinction.

4.29. Thus a program for purposes of scientific research will be designed to achieve a clearly identified outcome while avoiding unintended negative consequences that will put the population or stock being studied at risk. This should include the identification of potential problems before fieldwork begins and monitoring of the risk of unintended negative consequences during empirical work.

IWC Criteria for Special Permit Whaling

4.30. The Scientific Committee of the IWC has spent many years considering how the broad concepts in the previous paragraphs apply to scientific research in the context of conservation and management of whales. Their most recent thinking is summarized in IWC (2009). According to IWC (2009), proposals for Special Permit research are to be structured according to:

a) Objectives of the study (Paras 4.9-4.13; 4.39a);

b) Methods to address the objectives (Paras 4.14-4.16; 4.39b)

c) Assessment of potential effects of catches on the stocks involved (Paras 4.27-4.29; 4.39d);

d) A note on the provisions for co-operative research for both field and analytical studies (Paras 4.17-4.26; 4.39c); and

e) A list of scientists proposed to be sent to intersessional review workshops (Paras 4.17-4.26; 4.39c).

4.31. According to IWC (2009), the objectives of the study should:

a) Be quantified to the extent possible;

b) Be arranged in two or three categories (primary, secondary, ancillary);

c) Include a statement for each primary category regarding whether it involves lethal sampling, non-lethal sampling, or both;

d) Include at least a brief statement of the value of each primary objective assessed according to the ability to i) improve the conservation and management of whales stocks; ii) improve the conservation and management of other living marine resources in the ecosystem; and/or iii) test hypotheses not directly related to the management of living marine resources; and

e) Refer, particularly for d(i) and d(ii), to past recommendations of the Scientific Committee, carrying out implementations or reviews of the RMP, improved understanding of other high priority issues, or recommendations of other inter-governmental agencies.

4.32. According to IWC (2009), the methods should include:

a) Field methods that describe the species studied, the number, time frame and area; the sampling protocol for lethal aspects; and an explanation of why non-lethal methods or analyses of past data are insufficient;

b) Laboratory methods;

c) Analytical methods, including when appropriate estimates of whether the proposed sample sizes will be sufficient to provide accurate answers to the questions being studied; and

d) A time frame with intermediary targets.

4.33. According to IWC (2009), the assessment of potential effects of the proposed take on the stock should include:

a) A summary of what is known concerning stock structure in the area concerned;

b) An estimate of abundance of the species to be studied, including an assessment of the level of uncertainty of the estimates of abundance;

c) Submission of a simulation study on the effects of permitted takes on the catch, taking into account uncertainty and projecting forward for the life of the proposed permit, and into the future.

4.34. IWC scientists, like marine mammal biologists in general, understand that sometimes lethal take can provide information that other means of study cannot (Paras 4.31, 4.32). For example, although progress is being made, there are still no effective non-lethal means of aging whales, so if age information is absolutely required, then lethal take is also required.

4.35. Lethal take destroys the object of study and thus eliminates the possibility of future information gained from the animal that is killed. Thus, scientists must ask how much information is gained using a lethal method relative to the information gained using a non-lethal method. Consequently, before using lethal take, one must carefully weigh the balance between the immediate information produced by killing the individual animal and the loss of future information that could be obtained were a non-lethal method used. In my opinion, only when the balance is strongly in favor of the former should the lethal take be used. That is to say, the information gained must be proportional to the impact resulting from the loss of the individual.

4.36. The Society for Marine Mammalogy, the only international professional society of marine mammalogists, recently published guidelines for treatment of marine mammals in field research in its official journal *Marine Mammal Science.* These guidelines recognize that lethal take may sometimes be appropriate and state that (Gales et al 2009, pg 736):

a) researchers should use alternative non-lethal procedures when they are available and satisfy the objectives of the research;

b) animals should be killed in the most humane and rapid method available;

c) any population or stock-scale impacts should be minimized through prudent selection of animals (e.g., avoidance of reproductive females if possible) and sample size; and

d) where possible on-going activities outside the research community (e.g., hunts, by-catch events, strandings) should be utilized as a source of material for scientific studies of marine mammals.

4.37. The IWC criteria also recognize that when a scientific study is motivated by an important applied problem such as the conservation and management of whales, another crucial dimension is whether the knowledge extracted from the data can be used to answer the motivating applied problem. If the work cannot provide an answer to the motivating problem, it has failed in the key aspect of scientific inquiry, even if it produces other data. That is, a program that is motivated by an applied problem such as the conservation and management of whales must contribute to knowledge that informs the motivating problem. It is the responsibility of the proposers to demonstrate the objectives are both achievable with the methods proposes and that the work will contribute to the motivating applied problem.

Assessment Criteria Used in This Report

4.38. Consistent with the criteria for generally accepted scientific research and the IWC criteria described above, I now describe what I consider to be the essential characteristics of a program for purposes of scientific research in the context of the conservation and management of whales.

4.39. A program for the purposes of scientific research in the context of conservation and management of whales:

a) has defined and achievable objectives that aim to contribute knowledge that is important to the conservation and management of whale stocks;

b) uses appropriate methods that are likely to achieve the stated objectives, including:

   (i) lethal methods only where the objectives of the research cannot be achieved by any other means (for example, by the analysis of existing data and/or the use of non-lethal research techniques);

   (ii) setting sample sizes using accepted statistical methodology; and

   (iii) linking mathematical and statistical models to data consistently;

c) includes periodic review of research proposals and results and adjustment in response to such review; and

d) is designed to avoid adverse effects on the stocks being studied.

5. DESCRIPTION AND ASSESSMENT OF JARPA AND JARPA II AS PROGRAMS FOR PURPOSES OF SCIENTIFIC RESEARCH IN THE CONEXT OF CONSERVATION AND MANAGEMENT OF WHALES

5.1. It is now possible to provide an overview of the relevant aspects of JARPA II together with an assessment of those aspects against the essential characteristics of a program for purposes of scientific research in the context of conservation and management of whales. Although my primary focus is JARPA II, several aspects of the assessment are retrospective (e.g. peer review) and others are prospective (the feasibility of the research plan to achieve the goals). Consequently, I consider both JARPA and JARPA II.

5.2. In brief, the analysis in this section leads to the following conclusions:

a) The objectives of JARPA II are broad and poorly defined, often based on science by assertion in which statements are made as if they have been demonstrated but they in fact have not, and are formulated in a way that conflates exploration and exploitation.

b) Although a variety of empirical methods are used, the majority of effort in JARPA II is directed toward lethal take, with sighting surveys compromised because they are conducted in conjunction with lethal take. The connection between JARPA II as a field activity and management models such as the RMP has not been demonstrated, and the process for setting sample sizes in JARPA II is not based on solid statistical reasoning or analyses of the accuracy required to meet the stated objectives.

c) Individuals participating in JARPA II are disconnected from the self-correcting community of scientists and have not demonstrated the ability to revise or correct their work or methodologies, in particular by changing their minds concerning lethal take. The majority of the work conducted in association with JARPA and JARPA II is published outside the standard peer-review process and much of the work that is published in standard peer-reviewed literature refers only to the physiology and biochemistry of reproduction in whales, topics that are irrelevant to the stated objectives of JARPA and JARPA II.

d) There is no record of any attention being directed to avoiding unintended negative consequences in the design of JARPA II.

5.3. By reference to the conclusions in Para 5.2, the general practice of science, and the IWC criteria for Special Permit Research, I conclude that JARPA II is not a program for purposes of scientific research in the context of conservation and management of whales.

A program for purposes of scientific research in the context of conservation and management of whales has defined and achievable objectives that aim to contribute knowledge that is important to the conservation and management of whales

Vague & general objectives

5.4. Neither the goals of JARPA, nor those of JARPA II, have been clearly stated as defined and achievable objectives, nor as scientific questions or hypotheses that will contribute knowledge important to the conservation and management of whales.

5.5. In 1987 the objectives of JARPA were summarized as:

- Objective 1: Estimation of biological [demographic] parameters to improve the stock management of the Southern Hemisphere minke whale.
- Objective 2: Elucidation of the role of whales in the Antarctic marine ecosystem.

5.6. In 1995-97 two additional objectives were added:

- Objective 3: Elucidation of the effect of environmental change on cetaceans [whales and dolphins].
- Objective 4: Elucidation of the stock structure of the Southern Hemisphere minke whales to improve stock management.

5.7. Objective 1 was relevant to the NMP but is not relevant to the RMP; in addition, it was not achieved. Objectives 2, 3, and 4 are so broad that they can be used to justify almost any activity. Objectives 3 and 4 were added with little or no justification or connection to results that had previously been obtained under the program at the date of their addition. A program for purposes of science research will adjust its goals and objectives as information is obtained and analyzed, but this needs to be done with clear justification and reference to results obtained to date.

5.8. JARPA II continues the pattern established by JARPA of having broad objectives (IWC 2007a, pg 6):

- Objective 1: Monitoring of the Antarctic ecosystem.
- Objective 2: Modeling competition among whale species and developing future management objectives.

- Objective 3: Elucidation of temporal and spatial changes in stock structure.
- Objective 4: Improving the management procedure for minke whale stocks.

5.9. The objectives of JARPA II comprise a mixture of ecological monitoring and modeling (Objectives 1 and 2), field work (Objective 3), and management (Objective 4) with little, if any, intellectual connection. These objectives demonstrate confusion between monitoring (which may be important if tied to management, but cannot be considered research since there is no focused question or hypothesis) and management on the one hand, and alleged scientific investigation on the other.

5.10. Since lethal take without demonstrated scientific need is involved, the objectives of JARPA II blur potential scientific exploration and resource exploitation. Furthermore, as with JARPA, the objectives are so broad as to allow almost any activity.

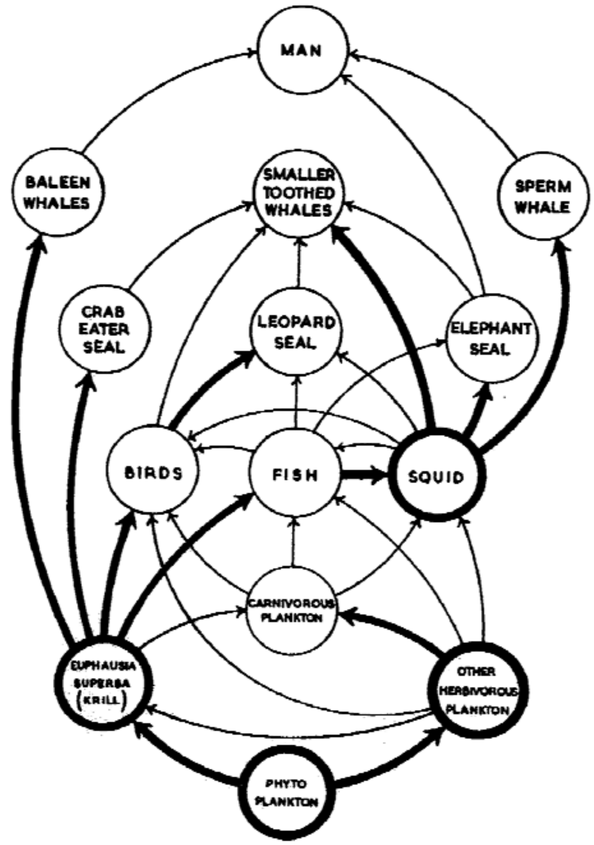
The ‘krill surplus’ hypothesis

5.11. As described above (Para 4.9ff), the application of an overarching conceptual framework should lead to a set of focused questions or hypotheses to be investigated since without clear questions the likelihood of developing new knowledge is slight.

5.12. However, the only clearly identifiable hypothesis in JARPA or JARPA II is the krill surplus hypothesis, according to which the over-harvesting of the great whales lead to a krill surplus, which in turn lead to an abundance of minke whales. For example, Tamura and Konishi (2009) wrote “[t]his rapid decreasing of large baleen whales species provided the annual surplus of krill as much as 150 million tonnes (Laws, 1977ab). This surplus became available for other krill predators…This phenomenon is called ‘krill surplus from the depletion of baleen whales’” (pg 23).

5.13. Among JARPA workers, the krill surplus hypothesis quickly went from hypothesis (that is, an idea to be investigated and possibly rejected) to theorem (that is, a demonstrated result whose truthfulness is known). For example in the review of JARPA undertaken by the government of Japan, Murase et al (2006) wrote “‘[k]rill surplus’ caused by intensive commercial harvesting of large whales…has been central theorem of the Antarctic ecosystem study” (pg 1). In describing the possible 'krill surplus' as a 'central theorem', Murase et al suggest that it has already been proven. This is simply not the case, as will be explained below.

5.14. The Antarctic continues to be unveiled as an ecosystem of intriguing complexity in which simple predictions fail (Karentz and Bosch 2001). The figure below (Mackintosh 1965) illustrates the network of interactions between predators and prey (arrows going from prey to predator).



5.15. From this figure, it is clear that the krill surplus hypothesis as applied in JARPA and JARPA II deals with only a small part of the entire Southern Ocean ecosystem. In addition, neither JARPA nor JARPA II are capable of testing it (Paras 5.36-5.37).

Data collected and the RMP

5.16. The literature concerning JARPA and JARPA II contains a variety of references to whaling policy that will be based on Maximum Sustainable Yield (MSY), which, as described above, has effectively been discarded by the IWC (Para 3.23ff). As noted, the RMP explicitly avoids the use of lethal-take data as a means of estimating abundance, instead placing a strong emphasis on data obtained by means of sighting surveys.

5.17. JARPA and JARPA II provide no demonstration of how the fieldwork undertaken in those programs would actually contribute to the analysis of MSY, MSYR, or to improving flaws in the RMP. In particular, JARPA II does not make clear how the improvement of management procedures for minke whale stocks can be considered scientific research, which might be appropriate if the RMP had been shown to be seriously flawed. However, neither JARPA nor JARPA II has demonstrated the existence of serious problems with the RMP.

5.18. JARPA was not relevant to the RMP, which (unlike the NMP) deliberately does not depend on accurate estimates of demographic parameters. In spite of this, JARPA II continues along the same path as JARPA. In particular, the collection of demographic parameters of whales by lethal take remains central in JARPA II, but has no relevance to the RMP.

Ecosystem model

5.19. At the meeting of the IWC that followed the 2006 Intersessional Workshop (IWC 2007b), "Japan re-iterated the goal of JARPA II, i.e. to develop an ecosystem model leading to sustainable use through multi-species management" (IWC 2007b, pg 41). Ecosystem-level models refer to conceptual, mathematical, or statistical models that include many components of the ecosystem, rather than a focus on a single species.

5.20. Although the development of ecosystem-level models is a foundation for Ecosystem Based Fisheries Management (Mangel 2010a), the contribution of JARPA or JARPA II as field programs to ecosystem-level management models is never made clear. Even though JARPA II's objectives have changed, its practice has not been altered so as to collect the type of data required for a far broader ecological study (see Paras 5.36-5.37).

5.21. One of the justifications of JARPA and JARPA II is that they will provide the scientific information that is required for the resumption of commercial whaling. It is remarkable that the JARPA and JARPA II documents lack even the beginnings of a bioeconomic model providing investigation about the required biological and economic circumstances to make commercial whaling on minke whales feasible, although it had been recognized long before JARPA began that a bioeconomic model would provide key insights into the future commercial whaling of minke whales (Lockyer 1976). Such models are clearly the province of scientific research since they provide the biological, economic and mathematical foundations for effective conservation and management as time-dependent phenomena (Clark 2010).

5.22. In summary,

- It is difficult to impossible to clearly identify the hypotheses of either JARPA or JARPA II;

- Both programs offer broad objectives that conflate science, management, and exploitation;
- Their stated objectives could be used to justify almost any activity that Japan wished to pursue.
- Their contribution to management remains undemonstrated and the potential of JARPA II to bring new knowledge about the conservation and management of whales is very low, if it indeed exists at all.

In my opinion, JARPA II fails to meet the essential first characteristic of a program for the purposes of scientific research in the context of conservation and management of whales.

A program for the purposes of scientific research in the context of conservation and management of whales employs appropriate methods likely to achieve the stated objectives, including (a) use of lethal methods only where the objectives of the research cannot be achieved by any other means (i.e. by the analysis of existing data and/or the use of non-lethal research techniques); (b) setting sample sizes using accepted statistical methodology; and (c) linking mathematical and statistical models to data consistently

Appropriate empirical tools

5.23. Scientific research on whales in the Southern Ocean can use a variety of tools for empirical research including:

a) sighting surveys in which whales are counted from ships or aircraft (including photo-identification of individuals);
b) lethal take;
c) DNA analyses based on biopsies;
d) biochemical analyses; and
e) satellite tagging.

Sightings surveys

5.24. Sightings surveys are a common feature of research in all global whale populations and if conducted appropriately may be a useful empirical tool for assessing the abundance and distribution of whales. Recent JARPA II cruises sighted blue, fin, sei, minke, humpback, southern right, sperm, and southern bottlenose whales (e.g. Ishikawa et al 2008).

5.25. Sighting surveys can provide information on population density (Burt and Borchers 1997), movement (Bannister et al 1999, Rock et al 2006), the relationship between physical habitat and whale distribution (Kasamatsu et al 2000) and the relationship between the abundance of food (krill) and whale distribution (Murase et al 2002).

5.26. As described above (Para 3.23ff), the RMP uses sighting surveys in order to estimate abundance and does not rely on lethally acquired information. However, some of the sighting surveys in JARPA and JARPA II are compromised because their methods involve both counting whales and preparation for lethal take.

Lethal take

5.27. In contrast to sighting surveys, lethal take is not a common feature of research in all global whale populations. While it cannot be excluded that there may be situations in which lethal take may contribute to a program for purposes of scientific research in the context of conservation and management of whales, JARPA and JARPA II simply assert but do not demonstrate that lethal take is required. In addition, lethal methodology is a disproportionate focus in JARPA and JARPA II.

5.28. Japan sought to justify lethal take as a means of obtaining age estimates that could then inform the rate of natural mortality (required for the NMP but not the RMP), but, as noted in the final review of JARPA, the effort failed.

5.29. This is because there are significant problems with the lethally derived data used for aging. Ear plugs of whales have a structure of alternating light and dark bands. Thus, in principle the age of a whale can be determined by counting the bands, much as with tree rings (Morris 1972, Roe 1967, Lockyer 1974, de la Mare 1985). However, the difficulties in the interpretation of growth layers make ear plug growth layers only somewhat reliable indicators of age. Furthermore, there are problems with reading the ear plugs at all and often a large number of the killed animals do not provide readable ear-plugs (Lockyer 2010).

5.30. As described in Para 4.14 a tool should only be selected for use after evaluating its effectiveness in achieving the stated objectives. Japan conducted no such evaluation. For ear plugs such an evaluation was done only after nearly 25 years of JARPA and JARPA II (Lockyer 2010) and ear plugs failed to provide information about the age dependence of the rate of natural mortality. Whether alternatives exist or not for aging, the approach of JARPA had demonstrably failed, but JARPA II continues along this track.

Other tools

5.31. Other common tools used in the study of populations of whales include

a) DNA analysis based on biopsies;

b) biochemical analyses; and

c) satellite tagging.

5.32. In the last 20 years, DNA and other molecular technologies for population studies have advanced enormously. A small sample of tissue now yields a sufficient amount of DNA for many different types of analyses on population structure, animal gender, inter-relatedness of individuals and other population scale parameters (Waples and Gaggiotti 2006).

5.33. Similarly, it is now possible to measure the concentration of many pollutants in whales by taking non-lethal biopsy samples (Kunito et al 2002) and to assess reproductive status from hormone concentrations in the blubber of minke whales (Mansour et al 2002). Awruch et al (2008) demonstrated that size at maturity in a shark could be obtained using blood samples measuring hormones. These papers suggest a promising avenue of investigation for a non-lethal method of determining reproductive status in whales.

5.34. Over the last 20 years, tagging whales with radio transmitters and associated technology for collecting and analyzing data has progressed rapidly (Fedak 2004, Freitas et al 2008). Satellite tags have an antenna that protrudes through the skin, so that the whale eventually rejects them much like a splinter is ejected. Mate et al (2007) reviewed the advances in satellite tags. Currently, tags last long enough to cover either leg of the annual migration or the whole feeding season and within the decade will likely last for multiple seasons. These longevities are sufficient to answer the critical questions about stock structure required to apply the RMP multi-stock rules.

5.35. In 1987, the Government of Japan (Japan 1987, pg 43) noted that "[i]f mark [or tag] and mark recapture could be available both in the low latitude (breeding ground) and the high latitude (feeding ground), this method [mark-recapture] would certainly produce information with the highest accuracy ever obtained by any other methods ever adopted in the past for ascertainment of stock movement, migration, and identification". At the time that this was written, the longevity of tags was only about 3 weeks. However, this gold standard of methodology - called for by Japan nearly 25 years ago - can now be achieved. That is, non-lethal means are now a practicable way for determining stock structure.

Linking methods to objectives

5.36. Japan has suggested that JARPA and JARPA II can test the krill surplus hypothesis. (Japan 2000, pg 1). However, neither JARPA nor JARPA II is sufficiently broad or deep to be able to test the krill surplus hypothesis as a scientific hypothesis. Indeed, it may be impossible to test the krill surplus hypothesis at all (Ainley et al 2007). Nicol et al (2007) observed

> It [the krill surplus hypothesis] is just difficult to support or refute without appropriate long-term, systematically collected, dataset on krill and its major predators. With a few notable exceptions, we are not in a position to be able to indicate whether most of the major krill consumers have globally increased or decreased as a result of the demise of the great whales, nor how these predators might now be responding to the recovery of some of these whale populations. Furthermore, we remain unable to estimate robustly global krill consumption now or in the past; data which are essential for examining the krill surplus hypothesis (pg 292).

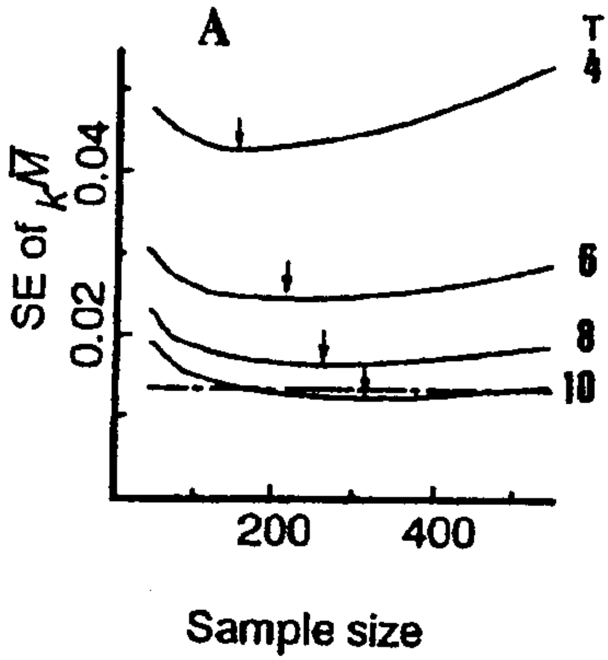
5.37. That is, given the enormous difficulties and the many uncertainties surrounding the krill surplus hypothesis, a program intended to investigate it would need to begin with a broad focus on the interactions between different predators of krill and krill consumption by all such predators (see image in Para 5.14). JARPA and JARPA II do not do this since their narrow focus is purportedly on three (in practice, essentially just one) species of whale. Rather, JARPA and JARPA II have used the krill surplus hypothesis to conflate research and exploitation.

Setting sample sizes

5.38. In the case of both JARPA and JARPA II, sample size indicates the number of whales to be killed. It is very difficult to understand the statistical basis for setting the level of lethal take in either JARPA or JARPA II.

5.39. Early in JARPA, Tanaka et al (1992) computed the sample size (lethal take) associated with the error in an estimate of the rate of natural mortality. The resulting curves (Tanaka et al 1992, Figure 6) were very flat. I have reproduced one panel below (the others are similar).

5.40. The x-axis in this figure is the sample size (the number of whales killed) and the y-axis is a measure of error in the estimate of the rate of natural mortality. The different curves in this figure show the error in the estimate of the rate of natural mortality plotted as a function of the sample size under various assumptions about the accuracy of the data.



5.41. The arrows denote the sample size that gives the minimum standard error. All else being equal, these arrows would point to the sample size that one would choose were one's goal to minimize the error of the estimate.

5.42. However, the curves are very flat, suggesting that many fewer whales could be taken with only a minimal loss of accuracy. For example, using the curve marked T=10, note that taking 100 whales rather than 300 whales only marginally decreases the accuracy of the estimate of natural mortality rate, suggesting that many fewer whales could be taken without compromising the resulting analysis. That is to say, many fewer whales killed will produce virtually the same level of accuracy.

5.43. Indeed, Tanaka et al (1992) themselves noted that takes in the range of 200-400 whales provided the same accuracy, but then stated that "[h]owever, in the actual research, other factors should be taken into consideration" (pg 419) to increase sample size. The other factors are not explained in the context of objectives nor are they justified through any statistical considerations. There is no analysis provided to show how either knowledge or management would be improved by having the marginally improved estimate of the rate of natural mortality associated with a take of 300 rather than 100 individuals.

5.44. Lack of statistical clarity continues in JARPA II. For example, in responding to the discussion of the proposal for JARPA II by Childerhouse et al (2006), Hatanaka et al (2006) wrote that catches "under JARPA II have been calculated as the minimum required to obtain statistically significant data. *Given that the stocks to be sampled are abundant and, for humpback and fin whales,* *increasing rapidly, it is quite logical that the sample size is correspondingly large*" (italics added).

5.45. This conclusion is not logical at all. According to generally accepted scientific and statistical methodology, the determination of a sample size must be grounded in statistical reasoning. Whether the stocks are sufficiently abundant may affect the practicability of taking a particular sample size, but it should in no way affect the actual determination of the sample size.

5.46. I consider that the spatial distribution of the lethal take is also important. The IWC has divided the Southern Ocean into six sectors for reporting catches and other data. In the years between the 1963/64 season and the 1985/1986 season the vast majority of Japanese minke whaling take was in IWC Areas IV and V [which are the areas closest to Japan], with very few whales taken from Areas I and II [where much more fuel and time would be needed to operate] (Ohsumi 1979).

5.47. The proposal for JARPA (Japan 1987) noted that "very little information was obtained [from commercial activity] for Area I and Area II" (pg 8). One might therefore expect the focus of a program for purposes of scientific research to be on Areas I and II, in order to gain more information about those regions, but instead the focus in both JARPA and JARPA II is in Areas IV and V because it "makes the research more efficient" (pg 8). That may be true if one measures efficiency in terms of whales killed per effort, but less so if efficiency is measured in terms of new knowledge.

5.48. That is, because the effort in JARPA II is in regions in which Japan traditionally whaled, JARPA II is collecting data that in large part already exist from commercial whaling (before JARPA) and JARPA itself. The potential development of new knowledge in this situation is very low.

Linking mathematical models to data

5.49. Ecosystem models are one of the objectives of JARPA II, but the JARPA II proposal (Japan 2005, pg 11) discusses modeling competition among whale species and future management objectives with no reference to other components of the ecosystem. The models used in Appendix 9 of the proposal for JARPA II (pg 81-82) do not require the detailed information that JARPA II sets out to collect.

5.50. Indeed, neither JARPA nor JARPA II offer explanation for the assertion made under those programs that to obtain the necessary data for the models requires lethal take, nor do they offer an explanation or indication as to how those data are to be used in the models. After nearly 20 years of JARPA effort, Mori and Butterworth (2006) offered a "first step towards modeling the krill-predator dynamics of the Antarctic ecosystem". Their model (pg 225ff) does not require the data from lethal take that is purportedly essential under JARPA II.

5.51. In summary,

- Although a variety of empirical methods are in principle employed in JARPA II, a majority of effort is devoted to lethal take for which there are other, non-lethal methods that can provide nearly all of the same information.
- The lethal take data are not required for the RMP.
- There are problems with the lethally derived data and many animals are killed without providing any useable data.
- Other tools (DNA and biochemical analyses from skin biopsies, satellite tagging) can provide much the same information as that provided by lethal take.
- Japan has not demonstrated that its objective of developing an ecosystem model (Para 5.8) is attainable with the data it collects through lethal research.
- JARPA II is insufficiently broad to test the krill surplus hypothesis, which has been treated not as a hypothesis but as a proved theorem in most of JARPA and all of JARPA II.
- The reasoning that underlies the setting of sample sizes (the number of animals killed) or the distribution of sampling effort is vague, unclear or simply wrong at times.

In my opinion, JARPA II fails when measured against the second essential characteristic of a program for purposes of scientific research in the context of conservation and management of whales.

A program for purposes of scientific research has periodic review of research proposals and results and adjustment in response to those reviews.

5.52. In the development of a program for the purposes of scientific research in any applied context, the responsibility is on the proposers to demonstrate that the objectives are important and attainable with the methods proposed and will contribute to the applied problem. This should be done through peer review of proposals and resulting papers. Although the proposals for JARPA and JARPA II had some form of review within the Scientific Committee of the IWC, there is no evidence that they went through rigorous and anonymous peer-review by experts in the field or that the proposals were substantially changed in response to the comments obtained in review.

5.53. Workers involved in JARPA began, and those in JARPA II continue with and consistently defend the position that 'lethal take is required' (Ohsumi 1995) with no demonstration of ability to change their minds or respond to feedback when lethal take is discussed.

5.54. For example, in 1998 JARPA workers argued "Genetic analyses using DNA can be conducted using biopsy sampling. However, the number of samples required in studies on stock identification in the case of the southern minke whale is large, and consideration of sampling collection should be taken into account" (IWC 1998, pg 412). DNA technology has changed so much since 1998 that this is no longer the case (Para 5.32) but there has been no change in the position of workers in JARPA II to reflect this.

5.55. The vague justifications for setting sample sizes (Paras 5.39-5.48) and the justification of lethal takes as a means of cost recovery (Ohsumi 1995) are examples of assumptions and policies that come from outside the scientific sphere. However, workers involved in JARPA and JARPA II have not exposed these assumptions and policies.

5.56. In 2010, Japan submitted a list of the scientific contributions of JARPA and JARPA II (and the north Pacific equivalents) to the IWC (Japan 2010). This list shows 195 IWC Scientific Committee and other meeting documents and 107 peer-reviewed journal publications listed for JARPA and JARPA II.

5.57. I divided the roughly one-third (107 of 302) of the publications that were peer-reviewed into categories of management (including genetic methods for stock identification and humane killing, ecology (including environmental toxicology), evolution and population genetics, and reproductive physiology or lipid biochemistry. The papers on management and ecology are potentially relevant to the objectives of JARPA and JARPA II, those in evolution less so, and those in reproductive physiology or biochemistry not even mentioned in the objectives.

5.58. Of these peer-reviewed publications, slightly less than half (51 of 107) deal with management or ecology. That is, only about one-sixth (one-half of one-third, or around 15%) of the articles are peer-reviewed and potentially relevant to the broadly stated objectives. Only about one-fourth of the papers in management or ecology appear in the ecological literature outside of IWC publications. In short, 1/12 (one-quarter of one-third) of the publications dealing with whale ecology and management have appeared in literature outside IWC publications.

5.59. Nearly 40% (39 of 107) of the peer-reviewed articles relate to reproductive physiology, or lipid biochemistry, which could be viewed as representing an opportunistic use of samples obtained because of the use of lethal methods. However, it is not clear how the knowledge about *in vitro* fertilization of minke whale eggs with previously frozen sperm, attempts to inject sperm into frozen and then thawed eggs, or to mature minke whale eggs *in vitro* are even remotely relevant to the objectives of JARPA or JARPA II as set out in the proposals for those programs.

5.60. Japan has asserted (IWC 2007) "that for ethical reasons, many western scientific journals refuse to accept papers based on lethal studies of whales". Since there is no supporting information for this assertion, it is difficult to determine how frequently such refusals occur and if they are based on the purported ethical reasons or the objective quality of the submitted work.

5.61. I note, however, that the IWC's journal *The Journal of Cetacean Research and Management* and *Marine Mammal Science* (see Para 4.36) both publish papers based on lethal studies, as long as the work is of sufficient quality and was conducted legally. Some more general journals, such as *Animal Behavior* do not consider work based on lethal studies, but others, for example *Oecologia* or *Polar Biology*, do consider work based on lethal studies.

5.62. In summary,

- The review of proposals for JARPA and JARPA II has been weak and the response to reviews even weaker.
- Workers in JARPA and JARPA II have not demonstrated an ability to respond to criticism or to admit being wrong.
- Workers in JARPA and JARPA II have not exposed assumption and policies that come from outside sources.
- Only about 15% of the papers produced by JARPA and JARPA II appear in peer-reviewed literature and are relevant to the objectives as laid out in the proposals.

In my opinion, JARPA II fails when measured against the third essential characteristic of a program for purposes of scientific research in the context of conservation and management of whales.

A program for purposes of scientific research in the context of conservation and management of whales is designed to avoid adverse effects on the stocks being studied

5.63. Estimates of the number of minke whales in the Southern Ocean have fluctuated considerably and are still highly uncertain, but for the purposes of this paper one may consider the estimate to be of the order of magnitude of 300,000-500,000 individuals (Gambell 1999, Gillespie 2005). Minke whale takes from 1974 to 1984 were of the order of 5,000 per year, and takes during the eighteen years of JARPA averaged just over 435 individuals per year. Under JARPA II, the takes number around 550 animals per year.

5.64. It was assumed in JARPA, and continues to be assumed in JARPA II, that lethal takes will have no effect on the dynamics of the stock (Nakamura 1991, 1993; Nakamura et al 1993). The danger is that this is a self-fulfilling prophecy: when one analyzes data assuming that there is no effect of the catch on the dynamics of the stock, then one will be forced to draw that conclusion since it is built into the analysis itself. Put another way: the assumption of no effect of research takes on populations is a preconceived conclusion.

5.65. Furthermore, if there are multiple local populations within the sampling area of JARPA II, the possibility exists that takes are unevenly distributed across different local populations, which can lead to different impacts. While unlikely, some depletion of small populations could be occurring and JARPA II would not be able to measure or monitor such impacts.

5.66. Using the estimated overall population size may be misleading for another reason. JARPA and JARPA II sample minke whale schools that are typically 1-4 individuals, containing a mixture of mature and immature individuals (Kato et al 1989, Fujise et al 1993, Kasamatsu et al 1993, Nishiwaki et al 2005). In population biology there is a phenomenon known as the Allee effect (Courchamp et al 2008, Mangel et al 2010) in which once the size of the population becomes sufficiently small (for example through anthropogenic effects) the population continues to decline, even if the original reason for the decline is removed. There are many causes for of Allee effects (Courchamp et al 2008), one of which is the disruption of social structure as would happen by removing individuals from small schools. The importance of social structure in minke whale feeding schools is still uncertain, but there is no mention in any of the JARPA or JARPA II literature of Allee effects, and nor of any efforts made to confirm that the populations under consideration do not show Allee effects.

5.67. In summary,

- Japan has not shown that JARPA II will not adversely affect the stocks, instead, it simply assumes that this will be so;
- There may be a whole range of indirect effects on the populations that are not even considered in JARPA II.
- A well-designed program of research would recognize these possibilities and check for them, even if the likelihood of an adverse effect on the overall population were small.

Thus, I consider that JARPA II is inconsistent with the fourth characteristic of a program for purposes of scientific research in the context of conservation and management of whales.

6. CONCLUSION

6.1. A program for purposes of scientific research in the context of conservation and management of whales must do much more than simply collect data; the data must be capable of forming the basis of new knowledge. Indeed, methods that generate the most data often do not generate the most knowledge. JARPA is an example of an activity that collected data but which failed to generate additional knowledge. On the basis of the materials I have reviewed, I consider that JARPA II will continue as an activity for the collection of data but, similarly to JARPA, will contribute little new knowledge relevant to the conservation and management of whales.

6.2. Scientific research work should begin with a question as opposed to an answer, since retrofitting a problem to a solution is almost never a good approach. Most importantly, the collection of data should never begin until one knows how it will be analyzed and used. Both JARPA and JARPA II began with an answer that lethal take is required and without clear plans of how data were to be/or will be analyzed or used.

Defined and achievable objectives that aim to contribute knowledge that is important to the conservation and management of whales

6.3. In 2005 the objectives of JARPA II were:

a) monitoring of the Antarctic ecosystem (including whales, krill and the feeding ecology of whales, and the effects of contaminants of cetaceans, monitoring of cetacean habitat);

b) modeling competition among whale species (including constructing a model of competition among whale species and new management objectives including the restoration of the cetacean ecosystem);

c) elucidation of temporal and spatial changes in stock structure; and

d) improving the management procedure for Antarctic minke whale stocks.

6.4. These objectives are based on considerable science by assertion, in which claims are stated as if they were demonstrated through rigorous study but actually are not.

6.5. The objectives of JARPA II are extremely broad and lack focus. Experience with JARPA suggests that the broad and vague objectives of JARPA II effectively allow any activity, and are used to provide justification for lethal take.

6.6. The RMP of the IWC provides a practical and well-tested approach for the management of future Southern Ocean whaling. It is an excellent compromise between the complexity of the model and the availability of data, and is capable of dealing with the high levels of uncertainty in the Southern Ocean ecosystem.

6.7. The data that are proposed for collection during JARPA II are not required for the RMP and the information on stock mixing (which is today better collected through combinations of satellite tagging and genetic analysis) will only peripherally contribute to any reconsideration of IWC regulations concerning stocks. Thus, the potential applicability of JARPA II to the RMP is low, if it exists at all.

6.8. In my opinion, JARPA II fails to meet the first criterion to characterize it as program for purposes of scientific research in the context of conservation and management of whales.

Appropriate methods that are likely to achieve the stated objectives, including:

i. lethal methods only where the objectives of the research cannot be achieved by any other means (for example, by the analysis of existing data and/or the use of non-lethal research techniques);

ii. setting sample sizes using accepted statistical methodology; and

iii. linking mathematical and statistical models to data consistently;

6.9. The methodology of JARPA II includes modeling, sighting surveys, biopsies, and lethal take. However, the expressed requirement for lethal take is science by assertion and the contribution of JARPA II as a field program to management models is not demonstrated.

6.10. The mathematical models proposed in association with JARPA II are, to a very large extent, independent of the field data collected in JARPA II, especially the lethal data.

6.11. Sighting surveys, biopsies, and modeling are all effective empirical tools that are available to address the currently stated objectives. Lethal take is not required to meet the objectives of JARPA II.

6.12. Even if lethal take were required, the process for setting sample sizes of lethal take in JARPA II is not based on solid statistical reasoning or analyses of the accuracy required to meet objectives.

6.13. In my opinion, JARPA II fails to meet the second criterion to characterize it as program for purposes of scientific research in the context of the conservation and management of whales.

Periodic review of research proposals and results and adjustment in response to such review.

6.14. Most of the work conducted in association with JARPA and JARPA II is published outside the standard peer-review process. Much of the work that is published in standard peer-reviewed literature is on physiology and biochemistry of reproduction in whales, topics irrelevant to the stated objectives of JARPA and JARPA II. Only about 15% of the papers resulting from JARPA and JARPA II are both peer-reviewed and relevant to stated objectives.

6.15. Scientists in JARPA and JARPA II have demonstrated an unwillingness to change their minds, particularly with respect to the asserted requirement for lethal take.

6.16. In my opinion, JARPA II fails to meet the third criterion characterizing a program for purposes of scientific research in the context of conservation and management of whales.

Is designed to avoid adverse effects on the stocks being studied

6.17. There is no record that JARPA II is designed with any attention directed to avoiding unintended consequences.

6.18. In my opinion, JARPA II fails to meet the fourth criterion characterizing a program for purposes of scientific research in the context of conservation and management of whales.

Overall Assessment of JARPA II

6.19. JARPA II is an activity that collects data in the Southern Ocean. However, by reference to standard accepted practice of science and the IWC Special Permit criteria, it is not a program for purposes of scientific research in the context of conservation and management of whales.

7. LITERATURE CITED

Ainley, D.G., Clarke, E.D., Arrigo, K., Fraser, W.R., Kato, A., Barton, K.J., and P.R. Wilson. 2005. Decadal-scale changes in the climate and biota of the Pacific sector of the Southern Ocean, 1950s to the 1990s. *Antarctic Science* 17:171-182

Ainley, D., Ballard, G., Ackley, S., Blight, L.K., Eastman, J.T., Emslie, S.D., Lescroel, A. Olmastroni, S., Townsend, S.E., Tynan, C.T., Wilson, P. and E. Woehler. 2007. Paradigm lost, or is top-down forcing no longer significant in the Antarctic marine ecosystem? *Antarctic Science* 19:283-290

Awruch, C.A., Frusher, S.D., Pankhurst, N.W., and J.D. Stevens. 2008. Non-lethal assessment of reproductive characteristics for management and conservation of sharks. *Marine Ecology Progress Series* 355:277-285

Bannister, J.L. 2002. Baleen whales. Pg 62-72 in *Encyclopedia of Marine Mammals.* (W.F. Perin, B. Wursig, and J.G.M. Thewissen, editors). Academic Press, San Diego

Bannister, J.L., Pastene, L.A. and Burnell, S.R. 1999. First record of movement of a southern right whale *(Eubalaena australis)* between warm water breeding grounds and the Antarctic ocean, South of 60°S. *Marine Mammal Science* 15(4): 1337-1342

Branch, T.A., Matsuoka, K., and T. Miyashita. 2004. Evidence for increases in Antarctic blue whales based on Bayesian modeling. *Marine Mammal Science* 20:726-754

Burt, M and Borchers D.l. Minke whale abundance estimated from the 1991/92 and 1992/93 JARPA sighting surveys. Documents submitted to the 1997 interim JARPA review. SC/M97/21. SC/M97/23

Clapham, P.J. and C.S. Baker. 2002. Whaling, modern. Pg 1328-1332 in *Encyclopedia of Marine Mammals.* (W.F. Perin, B. Wursig, and J.G.M. Thewissen, editors). Academic Press, San Diego

Childerhouse, S., Gales, N., Baker, C. S., Bass, C., Berggren, P., Bickham, J., Breiwick, J., Brownell, R., Carlson, C., Charrassin, J-B., Cipriano, F., Clapham, P., Collins, T., Cooke, J., Cozzi, B., Dinter, W., Engel, M., Findlay, K., Fortuna, C., Funahashi, N., Gedamke, J., Groch, K., Iniguez, M., Kasuya, T., Kell, L., Kock, K-H., Krahn, M., Leaper, R., LeDuc, R., Mattila, D., Moore, S., Northridge, S., Olavarria, C., Palazzo, J., Panigada, S., Parsons, C., Perrin, W., Pomilla, C., Porter, L., Reijnders, P., Ridoux, V., Ritter, F., Robbins, J., Rogan, E., Rojas, L., Rose, N., Rosenbaum, H., Rowles, T., Sadler, L., Secchi, E., Senn, D., Simmonds, M., Sironi, M., Stachowitsch, M., Thiele, D., Urban, J., Wade, P., Van Waerebeek, K., Waples, R., Weinrich, M., Williams, R., Wilson, B., Zerbini, A. 2006. Comments on the government of Japan's proposal for a second phase of special permit whaling in Antarctic (JARPAII). *J. Cetacean Res. Manage.* 8 (Supplement): 260-261

Clark, C.W. 2006. *The Worldwide Crisis in Fisheries. Economic Models and Human Behavior.* Cambridge University Press, Cambridge, UK

Clark, C.W. 2010. *Mathematical Bioeconomics. The Mathematics of Conservation.* Third Edition. John Wiley & Sons, New York

Cooke, J. G. 1995. The International Whaling Commission's Revised Management Procedure as an example of a new approach to fishery management. Developments in Marine Biology 4 (*Whales, seals, fish and man* [AS Blix, L Walløe, and Ø. Ulltang, editors], Elsevier, Amsterdam): 646-657

Courchamp, F., Berek, L. and J. Gascoigne. 2008. *Allee Effects in Ecology and Conservation*. Oxford University Press, Oxford, UK

de la Mare, W.K. 1985. On the estimation of mortality rtes from whale age data, with paraticular reference to Minke Whales (*Balaenoptera acutorostrata*) in the Southern Hemisphere. *Rep. int. Whal. Commn* 35:239-250

de la Mare, W.K. 1986a. The sensitivity of MSY to the parameters of the Baleen whale model. *Rep. int. Whal. Commn* 36:425-427

de la Mare, W.K. 1986b. Fitting population models to time series of abundance data. *Rep. int. Whal. Commn* 36:399-418

de la Mare, W.K. 1986c. Simulation studies of management procedures. *Rep. int. Whal. Commn* 36:429-450

Donovan,G.P. 2002. International Whaling Commission. Pg 637-642 in Encyclopedia of Marine Mammals. (W.F. Perin, B. Wursig, and J.G.M. Thewissen, editors). Academic Press, San Diego

Fedak, M. Marine animals as platforms for oceanographic sampling: a 'win/win' situation for biology and operational oceanography. *Memoirs of the National Institute of Polar Research*, Special Issue, 58:133-147

Freitas, C., Lydersen, C., Fedak, M.A., and K.M. Kovacs. 2008. A simple new algorithm to filter marine mammal Argos locations. *Marine Mammal Science* 24:315-325

Fujise, Y., Ishikawa, H., Saino, S., Nagano, M., Ishii, K., Kawaguchi, S., Tanifuji, S., Kawashima, S. and Miyakoshi H. 1993. Cruise report of the 1991/92 Japanese research in Area IV under the special permit for Southern Hemisphere minke whales. *Rep. int. Whal. Commn* 43: 357-371

Gambell, R. 1999. The International Whaling Commission and the contemporary whaling debate. pg 179-198 in *Conservation and Management of Marine Mammals* (J.R. Twiss and R.R. Reeves, editors). Smithsonian Institution Press, Washington, DC

Gillespie, A. 2005. *Whaling Diplomacy.* Edward Elgar, Cheltenham UK & Northamption, MA, USA

Gopnik, A. 2009. *Angels and Ages*. Alfred A. Knopf, New York

Grinnell, F. *Everday Practice of Science. Where Intuition and Passion Meet Objectivity and Logic.* Oxford University Press, Oxford and New York.

Harrison, S., Quinn, J.F., Baughman, J.F., Murphy, D.D., and P.R. Ehrlich. 1991. Estimating the effects of scientific study on two butterfly populations. American

Naturalist 137:227-243

Hatanaka, H., Morishita, J., Goodman, D., Pastene, L.A., and Y.Fujise. 2006. Response to Appendix 2 [Childerhouse et al 2006]. *J. Cetacean Res. Manage.* 8 (Supplement): 262-264

Hilborn, R. and M. Mangel. 1997. *The Ecological Detective. Confronting Models with Data.* Princeton University Press, Princeton, NJ

Holt, S. 2004. Forward to the 2004 printing. Pg i-xxiii in R.J.H. Beverton and S.J. Holt. *On The Dynamics of Exploited Fish Populations.* Blackburn Press. Caldwell, NJ, USA

Ishikawa, H., Goto, M., Ogawa, T., Bando, T., Kiwada, H., Isoda, T., Kumagai, S., Mori, M., Tsunekawa, M., Ohsawa, T., Fukutome, T.,K., Koyanagi, T., Kandabashi, S., Kawabe, S., Sotomura, N., Matsukura, R., Kato, K., Matsumoto, A., Nakai, K., Hasegawa, M., Mori, T., Yoshioka, S., and T. Yoshida. 2008. Cruise Report of the Second Phase of the Japanese Whale Research Program under Special Permit in the Antarctic (JARPAII) in 2007/2008', Japan, 2000 *estimated*, IWC Doc. SC/60/O4

IWC. 1995. 1994 Report of the Scientific Committee, *Rep. int. Whal. Commn* 45, 1995, 80-85

IWC. 1998. Report of the Intersessional Working Group to Review Data and Results from Special Permit Research on Minke Whales in the Antarctic, Tokyo, 12-16 May 1997. (SC/49/Rep1):377-412

IWC. 2007. Report of the 58th Annual Meeting.

IWC 2007a. 'Report of the Intersessional Workshop to Review Data and Results from Special Permit Research on Minke Whales in the Antarctic, Tokyo 4-8 December 2006', IWC Doc. SC/59/REP1

IWC 2007b. Report of the 58th Annual Meeting, pp 6, 16-17, 39-45

IWC 2009. Process for the review of special permit proposals and research results from existing and completed permits. *J. Cetacean Res. Manage.* 11 (Suppl). 2009: 398-401

Japan. 1987. 'The Program for Research on the Southern Hemisphere Minke Whale and for Preliminary Research on the Marine Ecosystem in the Antarctic', Japan, March 1987, IWC Doc. SC/39/04

Japan. 1987a. 'The Research Plan for the Feasibility Study on "The Program for Research on the Southern Hemisphere Minke Whale and for Preliminary Research on the Marine Ecosystem in the Antarctic"', Japan, October 1987, IWC Doc. SC/D87/1

Japan 2000. 'The 2000/2001 Research Plan for the Japanese Whale Research Program Under Special Permit in the Antarctic (JARPA)', Japan, April 2000, IWC Doc. SC/52/O2

Japan. 2005. 'Plan for the Second Phase of the Japanese Whale Research Program under Special Permit in the Antarctic (JARPA II) – Monitoring of the Antarctic Ecosystem and Development of New Management Objectives for Whale Resources', Japan, 2005 *estimated*, IWC Doc. SC/57/01

Japan. 2010. Scientific contributions of JARPA/JARPA II and JARPN/JARPN II. IWC Doc. 62/20.

Jenkins, S.H. 2004. *How Science Works. Evaluating Evidence in Biology and Medicine.* Oxford University Press, New York

Karentz D, Bosch I. 2001. Influence of ozone-related increases in ultraviolet radiation on Antarctic marine organisms. *American Zoologist* 41:3-16.

Kasamatsu, F., Yamamoto, Y., Zenitani, R., Ishikawa, H., Ishibashi, T., Sato, H., Takashima, K. and Tanifuji, S. 1993. Report of the 1990/91 southern minke whale research cruise under scientific permit in Area V. *Rep. int. Whal. Commn* 43: 505-522

Kato, H., Hiroyama, H., Fujise, Y. and Ono, K. 1989. Preliminary report of the 1987/88 Japanese feasibility study of the special permit proposal for Southern Hemisphere minke whales. *Rep. int. Whal. Commn* 39: 235-248

Kirkwood, G.P. 1992. Background to the development of Revised Management Procedures. Reports of the International Whaling Commission 42:236-243

Kunito, T., Watanabe, I., Yasunaga, G., Fujise, Y. and Tanabe, S. 2002. Using trace elements in skin to discriminate the populations of minke whale in southern hemisphere. *Marine Environmental Research* 53: 175-1

Larkin, P. 1977. An epitaph for the concept of maximum sustained yield. *Transactions of the American Fisheries Society* 106:1-11

Laws, R.M. 1977a. Seals and whales of the Southern Ocean. *Philosophical Transactions of the Royal Society B* 279:81-96

Laws, R.M. 1977b. The significance of vertebrates in the Antarctic Marine Ecosystem. Pg 411-438 in *Adaptations Within Antarctic Ecosystems.* Proceedings of the 3rd SCAR Symposium on Antarctic Biollogy . G.A. Llano editor. Smithsonian Institution, Washington, DC.

Lockyer, C. 1974. Investigation of the ear plug of the southern sei whale, *Balaenoptera borealis*, as a valid means of determining age. *Journal of the International Council for the Exploration of the Seas* 36:71-81

Lockyer, C. 1976. Body weights of some species of large whales. *Journal of the International Council for the Exploration of the Seas* 36:259-273-81

Lockyer, C. 2010. Report of the Antarctic Minke whale ear plug experiment. IWC Doc. SC/62/IA11 (with an associated spreadsheet)

Ludwig, D. and C.J. Walters. 1985. Are age-structured models appropriate for catch-effort data? *Canadian Journal of Fisheries and Aquatic Sciences* 42:1066-1072

Mackintosh, N.A. 1965. *The Stocks of Whales.* Fishing News Books, London

Mangel, M. 2010. Scientific inference and experiment in Ecosystem Based Fishery Management, with application to Steller sea lions in the Bering Sea and Western Gulf of Alaska. *Marine Policy* 34:836-843

Mangel, M. 2010a. Review of C.W. Fowler *Systemic Management: Sustainable Human Interactions with Ecosystems and the Biosphere* Marine Mammal Science 26:757-760

Mangel, M., Marinovic, B., Pomeroy, C., and D. Croll. 2002. Requiem for Ricker: Unpacking MSY. *Bulletin of Marine Science* 70: 763-781

Mangel, M., Brodziak J.K.T., and G. DiNardo, 2010. Accounting for Variation in Mortality and Allee Effects When Computing Steepness for Strategic Fisheries Management. UCSC-SOE-10-33. Available at http://www.soe.ucsc.edu/research/report?ID=1596

Mansour, A.A.H., McKay, D.W., Lien, J., Orr, J.C., Banoub, J.H. Øien, N. and S. Stention2002. Determiniation of pregnancy status from blubber samples in minke whales (*Balaenoptera acutorostrata*) *Marine Mammal Science* 18:112-120

Mate, B., Mesecar, R., and B. Lagerquist. 2007. The evolution of satellite-monitored radio tags for large whales: One laboratory's experience. Deep-Sea Research II 54:224-227

Mori, M. and Butterworth, D.S. 2006. A first step towards Modeling the krill-predator dynamics of the Antarctic ecosystem. *CCAMLR Science* 13: 217-277

Morris, P. 1972, A review of mammalian age determination methods. *Mammal* Review 2:69-104

Murase, H., Matsuoka, K., Ichii, T. and Nishiwaki, S. 2002. Relationship between the distribution of euphausiids and baleen whales in the Antarctic (35°E-145°W). *Polar Biol* 25: 135-145. *document are available on the Institute of Cetacean Research website* (http://www.icrwhale.org/JARPAReview2.htm)

Murase, H., Nishiwaki, S., Ishikawa, H., Kiwada, H., Yoshida, T. and Ito, S., 2006. Results of the cetacean prey survey using echo sounder in JARPA from 1998/99 to 2004/2005. IWC Doc. SC/D06/J21

Nakamura, T. 1991. A new look at a Bayesian cohort model for time-series data obtained from research takes of whales. *Rep. int. Whal. Commn* 41: 345-348

Nakamura, T. 1993. Two-stage Bayesian cohort model for time-series data to reduce bias in the estimate of mean natural mortality rate. *Rep. int. Whal. Commn* 43: 343-348

Nicol, S., Croxall, J., Trathan, P., Gales, N., and E. Murphy. 2007. Paradigm misplaced? Antarctic marine ecosystems are affected by climate change as well as biological processes and harvesting. *Antarctic Science* 19:291-295

Nishiwaki, S., Ishikawa, H. and Fujise, Y., 2005. 'Review of general methodology and survey procedure under the JARPA', Japan, 2005 estimated, Government of Japan document number JA/J05/JR2. *available on the Institute of Cetacean Research website* (http://www.icrwhale.org/JARPAReview2.htm

Ohsumi, S. 1979. Population assessment of the Antarctic minke whale. *Rep. int. Whal. Commn* 29:407-420

Ohsumi, S. 1995. The Necessity of Employing Lethal Methods in the Study of Whale Resources. From 'Research on Whales, ICR, 1995).   Available at http://luna.pos.to/whale/icr_rw_oh.html

Ohsumi, S. 1997. Development of objectives in the JARPA. Paper in the submitted in the interim review of JARPA

Pickett, S.T.A., Kolasa, J., and C.G., Jones. 2007. *Ecological Understanding. The Nature of Theory and the Theory of Nature.* Academic Press/Elsevier, Amsterdam

Polanyi, M. 1969. The Republic of Science: Its political and economic theory. Pg 52-72 in *Knowing and Being* (M. Greene, editor). University of Chicago Press, Chicago, IL

Raymo, C. 1991. *The Virgin and the Mousetrap. Essays in Search of the Soul of Science.* Viking, New York

Resnik, D.B. 2011. A troubled tradition. *American Scientist* 99:24-27

Ricker, W.E. 1975. Computation and Interpretation of Biological Statistics of Fish Populations. Bulletin 191, Fisheries Research Board of Canada

Roe, H.S.J. 1967. Seasonal formation of laminae in the ear plug of the fin whale. Discovery Reports 35:1-30

Rock, J., Pastene, L.A., Kaufman, G., Forestell, P., Matsuoka, K. and Allen, J. 2006. A note on East Australia Group V Stock humpback whale movement between feeding and breeding areas based on photo-identification. *J. Cetacean Res. Manage.* 8(3): 301-305

Tamura, T. and Konishi, K 2009. Feeding Habits and Prey Consumption of Antarctic Minke Whale (*Balaenoptera bonaerensis*) in the Southern Ocean. *Journal of Northwest Atlantic Fishery Science* 42: 13-25

Tanaka, S., Kasamatsu, F. and Fujise, Y. 1992. Likely precision of estimates of natural mortality rates from Japanese research data for Southern Hemisphere minke whales. *Rep. int. Whal. Commn* 42: 413-420

Tønnessen, J.N. and A.O. Johnsen. 1982. *The History of Modern Whaling*. University of California Press, Berkeley and Los Angeles, CA

Valiela, I. 2001. *Doing Science. Design, analysis, and communication of scientific research*. Oxford University Press, New York

Waples, R.S. and O.Gaggiotti. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* 15:1419-1439

Wiedenmann, J., Creswell, K., and M. Mangel. 2008. Temperature-dependent growth of Antarctic krill: predictions for a changing climate from a cohort model. *Marine Ecology Progress Series* 358:191-202

8. APPENDICES

Appendix A. Biographical Details Concerning Marc Mangel

(Web page: http://www.soe.ucsc.edu/~msmangel/)

Marc Mangel is Distinguished Professor of Applied Mathematics and Statistics, Jack Baskin Endowed Professor of Technology and Information Management, and Director of the Center for Stock Assessment Research at the University of California Santa Cruz, where he has served since 1996. At Santa Cruz, he also directed the Geographic Information Systems Laboratory (1996-1999), served as Associate Vice Chancellor, Planning and Programs (1997-1999) and chaired the Department of Applied Mathematics and Statistics (2007-09). Since 1 July 2010, he has also chaired the Program in Technology and Information Management.

From 1980-1996, Mangel was at the University of California Davis, where he served as Assistant, Associate and Full Professor for eight years in the Department of Mathematics and eight years in the Department of Zoology/Section of Evolution and Ecology. He chaired the Department of Mathematics (1984-1989) and was founding Director of the Center for Population Biology (1989-1993).

His awards include the Koopman Paper Prize from the Operations Research Society of America, 1982; JASA Applications Paper from the American Statistical Association, 1983; Joseph Myerhoff Fellowship, Weizmann Institute of Science, 1987; John Simon Guggenheim Memorial Fellowship, 1987; Fulbright Senior Research Fellowship, Oxford University, 1988; George Gund Foundation Distinguished Environmental Scholar, Case Western Reserve University,1992; Distinguished Statistical Ecologist, International Association for Ecology, 1998; Mote Eminent Scholar, Florida State University, 2000; Fellow, California Academy of Sciences, 2000; Fellow American Association for the Advancement of Science, 2003; UCSC Academic Senate Excellence in Teaching Award, 2003; Frohlich Fellow, CSIRO Hobart, 2006; Astor Lecturer, University of Oxford, 2007; Kaeser Lecturer University of Wisconsin, 2008; Fellow of the Royal Society of Edinburgh, 2009; the award for the best paper (out of 95) published in *The Transactions of the American Fisheries Society* for 2009, for their work on life history models of steelhead trout on the Central Coast of California, and Lamberson Ecology Trust Lecturer Humboldt State University, 2010.

Mangel has numerous journal publications and books that include *Decision and Control in Uncertain Resource Systems* (1985, Academic), *Dynamic Modeling in Behavioral Ecolog*y (with Colin Clark, 1988, Princeton), *The Ecological Detective. Confronting models with data* (with Ray Hilborn, 1997, Princeton University Press), *Dynamic State Variable Models in Ecology: Methods and Applications* (with Colin Clark, 2000, Oxford University Press), and *The Theoretical Biologist's Toolbox. Quantitative methods for ecology and evolutionary biology* (2006, Cambridge, University Press). He edited *Classics of Theoretical Biology* (A Special Issue of the Bulletin of Mathematical Biology. Part I: Volume 52 Numbers 1,2. Part II: Volume 53, Numbers 1,2), *Sex Allocation and*

Sex Change: Experiments and Models (Lectures on Mathematics in the Life Sciences, Volume 22) and *Proceedings of the Second International Symposium on Krill* (Canadian Journal of Fisheries and Aquatic Sciences 57(Supplement 3). He has supervised more than 50 undergraduate research projects or senior theses, 20 PhD students and 28 post-doctoral colleagues.

Mangel and Douglas Butterworth were the first two invited experts to the Scientific Committee of the Commission for the Conservation of Marine Living Resources (CCAMLR) and he served on the US delegation to CCAMLR in 1991. His work on southern ocean krill has been supported by NOAA Fisheries (1994-97), the US National Science Foundation (1998-2002) and the Lenfest Ocean Program (2006-2010). Mangel served for six years (1990-1996) on the Committee of Scientific Advisors of the US Marine Mammal Commission and in that role lead the effort to update the Principles for the Conservation of Wild Living Resources (Mangel et al 1996). He served on the Special Committee on Seals for the Natural Environment Research Council of the UK from 2004-2011, chairing it from 2008-2011.

Appendix B. Terms of Reference provided by the Government of Australia

The focus of your report should be on the Second Phase of the Japanese Whale Research Program under Special Permit in the Antarctic (JARPA II). However, your report should draw on references to the First Phase of the Japanese Whale Research Program under Special Permit in the Antarctic (JARPA), where it is relevant.

Your report should reflect your honest belief as to the issues and questions posed.

Please address the following matters/questions in your report:

a) identify and outline the essential characteristics of a program undertaken for purposes of scientific research; and

b) provide a critical analysis of the objectives, methodologies and other features of JARPA II and, in so doing, assess whether JARPA II has the essential characteristics referred to in paragraph (a).

Appendix C - Background Material Provided by the Government of Australia

The Government of Australia provided the following material:

- The International Convention for the Regulation of Whaling, 1946;
- A range of IWC documents relating to special permit whaling, including
  - resolutions of the Commission concerning special permit whaling and JARPA and JARPA II from 1987 to 2007;
  - relevant extracts of the annual reports of the Commission and Scientific Committee from 1985 to 2009, including discussions on special permit whaling and the RMP;
  - reports of the IWC interim and final reviews of JARPA;
  - summary of special permits issued 1951 to 1987;
- A collection of documents prepared by the Government of Japan, including
  - research proposals in relation to JARPA and JARPA II, from 1987 to 2005;
  - cruise reports in relation to JARPA and JARPA II, from 1988 to 2010;
  - special permits issued in relation to JARPA and JARPA II;
  - report of the Government of Japan review meeting of JARPA, 18-20 January 2005;
  - documents submitted to the IWC interim and final reviews of JARPA; and
  - publications listed on the Institute of Cetacean Research website.

In addition, I was provided with a range of scientific papers and publications. Any paper cited appears in the list of Literature Cited.