Whales, science, and scientific whaling in the International Court of Justice

Marc Mangel*a,b,1

“Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064; and bDepartment of Biology, University of Bergen, Bergen 6020, Norway

Edited by Alan Hastings, University of California, Davis, CA, and approved September 29, 2016 (received for review April 20, 2016)

PNAS Early Edition | 1 of 5

any Contracting Government may grant to any of its nationals a special permission authorizing that national to kill, take and treat whales for purposes of scientific research subject to such restrictions as to number and subject to such other conditions as the Contracting Government thinks fit...the killing, taking, and treating of whales in accordance with the provisions of this Article shall be exempt from the operation of this Convention. (10)

Scientific research is not defined in the ICRW; there are historical reasons for this (4). Indeed, there are about 40 international environmental agreements that mention scientific research without defining it (11).

According to its preamble, the objective of the ICRW is “to provide for the proper conservation of whale stocks and thus make possible the orderly development of the whaling industry” (10). To do this, the ICRW created the intergovernmental IWC, which is charged, among other things, with regulating whaling in the Southern Ocean. The IWC has a Scientific Committee that meets annually and provides advice to the Commission. However, between 1946 and 1982, the IWC failed in proper conservation of whale stocks and the orderly development of the whaling industry (4). For example, the largest ever annual take of whales occurred around 1960–15 y after the ICRW—and three of the four peak whale catches occurred subsequent to the ICRW (12).

After a number of failed attempts at better management methods, members of the Commission decided on a complete commercial moratorium, effective 1986, to allow the rebuilding of stocks and the development of a Revised Management Procedure. The decision for a complete commercial moratorium—rather than on a stock/species basis—was controversial (1, 4, 5).

The Japanese Whale Research Program Under Special Permit in the Antarctic. Simultaneous with the start of the moratorium, Japan initiated JARPA (www.icrwhale.org/sc/JARPA.html) by authorizing the taking, via special permit under Article VIII, 400 minke whales per year by the Institute of Cetacean Research. In 1987 the objectives of JARPA were: (i) estimation of biological (demographic) parameters to improve stock management of the Southern hemisphere minke whale (which was relevant to one of the previous but by then discarded management procedures, the New Management Procedure) and (ii) elucidation of the role of whales in the Antarctic marine ecosystem. In 1995–1997, two additional objectives were added: (iii) elucidation of the effect of environmental change on cetaceans and (iv) elucidation of the stock structure of the Southern Hemisphere minke whales to improve stock management.

In 2005, a review of JARPA was initiated (completed in 2006), and the follow-up program, JARPA II, was proposed (January)
and begun (November). Under JARPA II, Japan issued permits to take up to 850 minke whales, 50 fin whales, and 50 humpback whales each year. The objectives of JARPA II overlapped considerably with those of JARPA: (i) monitoring of the Antarctic ecosystem, (ii) modeling competition among whale species and developing future management objectives, (iii) elucidation of temporal and spatial changes in stock structure, and (iv) improving the management procedure for minke whale stocks. Both programs involved a mixture of monitoring and management; unlike JARPA, JARPA II had no time horizon.

The Case in the ICJ

In May 2010, after nearly 25 y of trying by diplomatic means to convince Japan to stop whaling in the Southern Ocean, Australia initiated a case in the ICJ. The written pleadings, expert testimony, and responses, and transcripts of the oral hearings can be found online at www.icj-cij.org/docket/index.php?p1=1&format=pdf&p2=3&code=aj&case=148&k=64&kp3=0 by selecting the tabs “Written Proceedings,” “Oral Proceedings,” etc. Video of the oral hearings is available at United Nations TV (webtv.un.org/search?term=whaling).

In its Application to the Court (13), Australia asked that the Court to order Japan to: (i) “cease implementation of JARPA II”; (ii) “revoke any authorizations, permits or licences allowing the activities which are the subject of this application to be undertaken,” and (iii) “provide assurances and guarantees that it will not take any further action under the JARPA II or any similar program until such program has been brought into conformity with its obligations under international law.”

The ICJ is the principal judicial organ of the United Nations and all member states of the United Nations are automatically members of the Court’s statute. The Court’s role is to settle legal disputes submitted to it by States. The Court has 15 Judges, elected by the Security Council and General Assembly for 9-y terms. These Judges, according to Article 2 of the ICJ statute, are of “high moral character” and possess “the qualifications required in their respective countries for appointment to the highest judicial offices or are jurists/consults of recognized competence in international law” (14). If, as in this case, a nation does not have one of its citizens on the Court, it may nominate a judge ad hoc (Australia’s nomination of Professor Hilary Charlesworth was readily accepted). Many of the Judges are distinguished academics who understand the process of research and scholarship, albeit in a nonscientific field.

A Program for Purposes of Scientific Research. I was asked by the government of Australia to develop criteria for a program for purposes of scientific research in the context of the conservation and management of whales in the Southern Ocean, and to assess JARPA II against those criteria. To assure my independence as a witness, I was not informed of Australia’s legal strategy and only saw it unfold during oral proceedings.

Thus, my task involved considerations of science as a process, and communicating these ideas along with my assessment of JARPA II to the Judges. My full analysis is included as Appendix 2 to the Australian written submission (15). My determination was that a program for purposes of scientific research in the context of conservation and management of whales required: (i) a conceptual framework leading to testable predictions, which is almost a definition of modern science; (ii) a process for setting sample sizes of lethal take based on solid statistical reasoning and analyses of the accuracy required to meet objectives; (iii) regular peer-review of research proposals and results; and (iv) design to avoid adverse effects on the stocks being studied.

Assessing JARPA II against these criteria, I concluded that “JARPA II is an activity that collects data in the Southern Ocean. However, it is not a program for purposes of scientific research” (15).

In its written proceedings (see the Counter Memorial in ref. 16), Japan offered no alternative definition of science or of a program for purposes of scientific research.

The Judgment of the Court

The entire Judgment can be found at www.icj-cij.org/docket/files/148/18136.pdf (17); in what follows I refer to paragraph (Para) numbers when either quoting or paraphrasing the Judgment. In addition, when I refer to “the Court,” I mean the majority of the Court in the decision, dropping “the majority of” for simplicity. I first discuss the standard of review that the Court used for objectively assessing JARPA II, then its decision not to define science, and finally the Judgment itself.

The Standard of Review. The ICJ usually goes directly to the interpretation of the relevant treaty provisions. Thus, the Court could have asked “Do the special permits issued by Japan in connection with JARPA II comply with the requirements and conditions prescribed by the provisions of the ICRW, the Schedule that operationalizes the treaty, and Annex P (a document of Scientific Committee-IWC that establishes criteria and conditions for evaluating special permits)?” Instead, the Court adjudicated by analyzing JARPA II to determine if it was for purposes of scientific research. This required establishing a standard of review: the criteria for an objective analysis of JARPA II. The approach used by the Court appears to have been borrowed from dispute settlement in the World Trade Organization (WTO). In certain WTO disputes between two States, adjudicators are required to determine whether a risk assessment undertaken by a WTO member is supported by “coherent reasoning and respectable scientific evidence” (18, 19).

The Court explains its standard of review in paragraphs 62–69 and 88 but not why this route was chosen (17). By choosing to objectively assess JARPA II using the standard of review described below, the Court introduced an objective process into international environmental law that may have a lasting impact (20, 21). The Court focused on legal—rather than scientific—assessment of the issues, but at the same time meaningfully engaged with the scientific evidence as it related to a legal standard (22). Sir Geoffrey Palmer, former Prime Minister of New Zealand, considers that with the standard of review the Court developed a demanding and rigorous approach to multilateral international conventions (23).

In applying this standard of review, the Court considered research objectives paramount: “Moreover, an objective test of whether a programme is for purposes of scientific research does not turn on the intentions of individual government officials, but rather on whether the design and implementation of a programme are reasonable in relation to achieving the stated research objectives...The research objectives alone must be sufficient to justify the programme as designed and implemented” (Para 97 in 17).

The Decision Not to Define Science. To begin, the Court established an important principle that science is fundamentally objective: “...whether the killing, taking and harvesting of whales pursuant to a requested special permit is for purposes of scientific research cannot simply depend upon that State’s perception (Para 61 in ref. 17). Thus, the case offered the Court an opportunity to offer a definition of science to be used in international law, as the US Supreme Court did in its Daubert decision (24; available at https://supreme.justia.com/cases/federal/us/509/579/case.html). However, even though the Court elaborated on science in paragraphs 73–86 of ref. 17, it chose to neither accept my definition (Para 86 in ref. 17) nor to offer its own definition of science. For further discussion of this point, see Gogarty (25).

Even though the Court did not define science, the process of science is interwoven in the standard of review for comparing the objectives, design, and implementation of JARPA II based on the arguments the Court read and heard (Para 88 in ref. 17):
(i) decisions regarding the use of lethal methods, (ii) the scale and use of lethal sampling, (iii) methodology used to select sample sizes, (iv) comparison of sample size and actual take, (v) the program’s time frame, (vi) the program’s scientific output, and (vii) the degree of coordination with other research programs.

The Court was able to adjudicate the case without defining science by interpreting the relevant phrase in Article VIII as cumulative. First, the Court asked whether the lethal take in JARPA II was potentially scientific research. Second, it asked whether the lethal take was for purposes of scientific research (Para 67 in ref. 17). The Court considered that these are cumulative; to be within Article VIII, a program had to involve scientific research and be for purposes of scientific research broadly speaking.

The Court’s Objective Assessment. In this two-step process, the Court first concluded that “the JARPA II activities involving the lethal sampling of whales can broadly be characterized as ‘scientific research.’ There is no need therefore, in the context of this case, to examine generally the concept of ‘scientific research.’” (Para 127 in ref. 17).

For the second step, the Court considered both design and implementation of JARPA II. Concerning design (Paras 224–225 in ref. 17), the Court noted that the scale of lethal sampling of minke whales more than doubled from JARPA to JARPA II and that JARPA II added two new species. However, the objectives of JARPA and JARPA II overlapped considerably and that “to the extent that the objectives are different, the evidence does not reveal how those differences lead to the considerable increase in the scale of lethal sampling in the JARPA II Research Plan” (Para 225 in ref. 17).

Furthermore, the sample sizes of fin and humpback whales were too small to provide any relevant information. In addition, fin whales were sampled outside of the center of their distribution and the whaling vessels were too small to randomly sample individuals. Finally, the process to determine the sample size for minke whales lacked transparency, something that I raised in written and oral testimony.

Concerning implementation (Para 226 in ref. 17), the Court noted a number of problems with each of the three species in JARPA II. First, no humpback whales were taken at all and Japan offered political reasons for this. Second, the take of fin whales (18 over 6 y rather than 30 a year) was so small as to be useless. Third, except for one season, the take of minke whales was far lower than the annual target but there was no adjustment to the lethal sampling program.

Finally, Japan had used a 12-y time frame for setting fin and humpback targets but a 6-y time frame for setting minke targets (had a 12-y time frame been used for minke whales, the annual target would be lower). The Court concluded that the research objectives could not remain viable given the mixed time frame, the under-take of minke whales, the essential absence of take of fin whales, and the exclusion of humpback whales and that lack of adjustment suggested motivations other than purposes of scientific research. For example, if an ecosystem model can be done without lethal data from fin and humpback whales, why are lethal data needed for minke whales?

Earlier in the judgment, the Court observed in an understated manner “that the first research phase of JARPA II (2005–2006 to 2010-2011) has already been completed (see paragraph 119 above), but that Japan points to only two peer-reviewed papers that have resulted from JARPA II to date. These papers do not relate to the JARPA II objectives...In light of the fact that JARPA II has been going on since 2005 and has involved the killing of about 3,600 minke whales, the scientific output to date appears limited” (Para 219 in ref. 17).

The Judgment. The Judgment is final, without appeal, and binding on the Parties. The Court found:

Unanimously that it had jurisdiction to assess the case;

By 12 votes to 4 that (i) the special permits granted by Japan in connection with JARPA II did not fall within the provisions of Article VIII, (ii) Japan had not acted in conformity with its obligations under the commercial moratorium, (iii) Japan had not acted in conformity with its obligations under the factory ship moratorium, (iv) Japan had not acted in conformity with its obligations under the Southern Ocean Sanctuary; and

By 13 votes to 3 that Japan had complied with its obligations informing the IWC with regard to JARPA II. (The three dissenting Judges here were different from the four who dissented above).

The Court ordered Japan to revoke existing authorization and permits to kill, take, or treat whales under JARPA II and to not grant any other permits under Article VIII related to JARPA II.

This was a stunning judgment: “It is rare in the International Court of Justice for a state to claim victory, and that claim to be entirely true. More commonly, each side gets to walk out to the court’s steps and tell the press how their arguments were vindicated in one way or another. Yet Australia may now rightly assert that it has been victorious in its case against Japanese whaling in Antarctica” (26).

Following the Judgment, Japan immediately cancelled JARPA II and did not hunt whales in the Southern Ocean in 2014–2015. Japan developed a new program (27), with a target of about 300 minke whales and no fin or humpback whales, and returned to the Southern Ocean in 2015–2016. Like JARPA and JARPA II, this new program is controversial (28).

Implications of the Case

The case has implications for the IWC, ICRW, and scientific whaling, for international treaties in general, and most generally for the interaction of science and law (or scientists, lawyers, and judges, more properly).

The IWC, ICRW, and Scientific Whaling. There are implications for the treaty itself. In its Memorial, Australia argued (Paragraph 2.35 ff in ref. 17) that the ICRW had evolved from a treaty that emphasized take to one that emphasized preservation. That is, both take and conservation are mentioned in the preamble of the ICRW: “Recognizing the interest of the nations of the world in safeguarding for generations the great natural resources represented by the whale stocks...conclude a convention to provide for the proper conservation of stocks and thus make possible the orderly development of the whaling industry” (17). Japan argued that the treaty was never intended to prohibit whaling. Australia argued that international norms had changed to the point that conservation meant preservation. However, the Court concluded that “The objectives of the ICRW are further indicated in the final paragraph of the preamble...Amendments to the Schedule and recommendations by the IWC may put an emphasis on one or the other objective pursued by the Convention, but cannot alter its object and purpose” (Para 56 in 17).

The Court made it clear that resolving the preservationist/conservationist [sustainable use (29)] dichotomy is neither a matter of science (30) nor law:

...The Court is aware that members of the international community hold divergent views about the appropriate policy towards whales and whaling, but it is not for the Court to settle these differences. The Court’s task is only to ascertain whether the special permits granted in relation to JARPA II fall within the scope of Article VIII, paragraph 1, ICRW” (Para 69 in ref. 17).

In summary, the Judgment demonstrated that there are objective means for determining where an activity is for purposes of scientific research but does not prevent future scientific whaling (31). In addition, the Court was silent on the main dispute at the IWC between states that believe that whales should not be killed and those that support sustainable, regulated take (23).

Other International Treaties. The judgment is generally significant because the ICJ found that a State has to explain itself when
dealing with nonbinding recommendations adopted by a treaty organization (cf. 32). The standard of review and subsequent judgment can be said to have turned soft-law [in this case, the nonbinding resolutions of the IWC; more generally declarations, resolutions, recommendations, charters, or codes of practice (33)] into hard-law because Japan now has an obligation to respond to the IWC recommendations concerning the use of lethal methods (cf. 34). It is clear that with this Judgment, the Court added to its repertoire a precedent of judicial review of the exercise of discretionary power by a State participating in a multilateral convention or treaty (35).

**Science and Law.** Although the Court did not define science for international law, it “administered an efficient process for the production of scientific expert evidence and was unusually vigorous in its questioning of the scientific experts called by the parties during oral proceedings” (36). That is, the Court relied on the experts but did not give carte blanche acceptance of expert opinion. By doing so, it provided a number of general lessons about science and law.

By the end of the oral proceedings, the Judges themselves were asking questions about sample size and testable hypotheses. Through their questions and the written Judgment, dissents, and separate opinions, the Judges demonstrated understanding of the need for testable questions, appropriate sample size (e.g., Paras 160–198 in ref. 17), how models and data are connected, adjusting field work according to circumstances (e.g., Paras 201, 206–212 in ref. 17), and peer review broadly defined (e.g., Paras 84, 155, 156, 219 in ref. 17). That is, it is clear that the Judges understood the links between objectives of a scientific study, methods to achieve those objectives, and the importance of appropriate sample size (even if they are not expert with the tools for determining such sample size).

Thus, **Whaling** shows that it is possible to build capacity for Judges without formal scientific training to recognize the merits of science in decisions, rather than simply deferring to expert opinion. Judges can indeed understand the process of science, if not all of the details, and in this way combine “the rigor of the scientific community with the requirements of the courtroom” (37). Judges will have to make decisions involving science without deciding on aspects that are clearly not part of the judicial function (38), and judicial decisions thus require putting the scientific information into a form appropriate for legal application so that fair and well-reasoned decisions are the result (39). This raises the question of what expectations we can have for scientists interacting with law.

Clearly, scientists have values (40) but they need to be separated from analysis when serving as expert witnesses. The question of whether or not a particular population of animals should be taken is an ethical one; the question of whether a particular population could sustain a specified take and the ecosystem consequences of that take is a scientific one. As scientists, when we advocate in a court or tribunal, our most important contribution is advocating for the process of science.

Thus, when cases involve scientific evidence: First, scientists should be prepared to explain that even when the method of strong inference does not hold, which is true for many environmental situations (41–43), what hypotheses, questions, or testable predictions were investigated and to explain them in clear and simple language. We can emphasize that science does not consist of simply accumulating data (44). In his comparison of Abraham Lincoln and Charles Darwin, Gopnik noted “All 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. It is in the leap of the data, not the heap of the data, as Muhammad Ali might have put it, that the advance lies” (45).

Second, scientists must make clear how statistical hypotheses are different from scientific hypotheses (41), and that they are clearly articulated in nontechnical language. Because there is now an international society called “Bayes and the Law” (https://sites.google.com/site/bayeslegal/home), dedicated to improving the quality of expert analysis and presentation using Bayesian methods, we should be prepared to explain the difference between Bayesian and frequentist statistics.

If frequentist statistics are used and the null hypothesis is rejected, then scientists must make clear the connection between the null hypothesis and the question of interest, and that the size of the effect is estimated. Because the most common error in interpretation is to draw an inference from failure to reject a null hypothesis, if the null hypothesis is not rejected then scientists should report the power of the statistical test, because action based on hypothesis testing without consideration of the power of the test may be disastrous (46). If Bayesian statistics are used, then the hypotheses under consideration should be carefully explained in nontechnical language, the posterior probabilities given, and effect sizes estimated.

Scientists should have flow charts, illustrating the logic of the models, if not the details, and be prepared to report if the model predictions were simply confirmations of a priori assumptions or something new was learned from the models. We must make clear that if one model fits the existing data best and has proven ability to explain new data, we might have a very high degree of belief in it. At the same time, we must emphasize that modeling is iterative, and that the current best model will likely be replaced by another model in the future.

Scientists should expect to report the level of peer-review of the work, including: (i) who selected the peer-reviewers, (ii) the source of their compensation (if any), and (iii) whether the material has appeared or is destined to appear in peer-reviewed scientific journals. We should be prepared to respond to these expectations, even if not asked all of them all of the time.

To be sure, there will be instances in which adjudicators have to make sense of technical information and in which understanding the process of science at the higher level described here will be insufficient (47). In that situation, judges may choose to have expert scientists as consultants, as well as witnesses. Justice Breyer of the US Supreme Court called for such involvement nearly 20 y ago, saying that “As society becomes more dependent for its well-being upon scientifically complex technology, we find that this technology increasingly underlies legal issues of importance to all of us... A judge is not a scientist and a courtroom is not a scientific laboratory, [but judges] must aim for decisions that, roughly speaking, approximately reflect the scientific state of the art” (48).

Even if there are party-appointed experts, when there is a matter of complex and contradictory technical assessment, Court appointed experts can play a valuable role. Article 50 of the ICJ statute explicitly allows this (49). Similarly Article 289 of the Law of the Sea Convention allows experts to sit with the court or tribunal but not vote and Article 27 of the Permanent Court of Arbitration allows an arbitral tribunal to appoint one or more experts to report on specific issues (50). Such an expert must be both a highly accomplished scientist and a highly skilled explainer, because the job will be to communicate to the judges the technical issues in a nontechnical way. For example, such an expert will surely have to translate between scientific and legal uses of the same word. Foster noted that “The challenge in many cases is not to determine what constitutes good or reliable science, but to reach a well-informed view of the existing science and the boundaries of the relevant scientific knowledge sufficient to assess whether there has been compliance with applicable international legal obligations” (50).

**Conclusions**

Even though the Court used a somewhat unique two-step procedure in this case, its approach to arbitration on science by taking at face value the stated objectives of a program and proceeding with its analysis using criteria based on the information that it gleaned from written material, oral arguments, and expert
testimony is very general. This provides a model in which the adjudicator decides the case using experts to provide guidance on how to get to the decision, but the adjudicator decides. Fitzmaurice (1) expects that the Whaling case will set the new international standard for the use of experts in disputes that turn on scientific facts. We have many contributions to make toward improving the future of science and law.


ACKNOWLEDGMENTS. I thank two anonymous referees, the Associate editor, Paul Daltrop, Brendan Gogarty, Susan Mangel, and Melissa Monk for comments on a previous version of the manuscript; Holly Doremus, Cymie Payne, Bruce Leaman, Daniel Peat, and Geoffrey Wandesforde-Smith for discussions about science and law; and Malgosia Fitzmaurice and Dai Tamada for sharing proofs of their book with me in advance of publication. My work on the case (August 2010–July 2013) involved a consulting agreement with the office of the Attorney General of Australia.