

Michigan Journal of International Law

Volume 38 | Issue 2

2017

Reliability of Expert Evidence in International Disputes

Matthew W. Swinehart

United States Department of the Treasury

Follow this and additional works at: <https://repository.law.umich.edu/mjil>



Part of the [Courts Commons](#), [Evidence Commons](#), [International Law Commons](#), and the [Organizations Law Commons](#)

Recommended Citation

Matthew W. Swinehart, *Reliability of Expert Evidence in International Disputes*, 38 MICH. J. INT'L L. 287 (2017).

Available at: <https://repository.law.umich.edu/mjil/vol38/iss2/6>

This Article is brought to you for free and open access by the Michigan Journal of International Law at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Michigan Journal of International Law by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

ARTICLE

RELIABILITY OF EXPERT EVIDENCE IN INTERNATIONAL DISPUTES

Matthew W. Swinehart*

INTRODUCTION

Even as the legal community, politicians, and the public at large continue to scrutinize the legitimacy of international dispute resolution,¹ one of the most influential features to have emerged in modern international law—the use of expert witnesses—has gone largely overlooked. Apart from the broader debates about whether and to what extent states should subject policy choices to international law or how treaty partners should design mechanisms to resolve disputes over those choices,² stands the significant role that expert witnesses play in those disputes. Where the disputing parties appoint experts, the adjudicator often must referee a “battle of the experts” and choose between two competing views and outcomes. Where the adjudicator has appointed an expert, the struggle is to ensure that the ultimate decision making authority is not delegated to the expert. Many recent high-profile international legal disputes—including a dispute

* Matthew Swinehart is a legal advisor for trade and investment agreement negotiations at the United States Department of the Treasury. The views expressed here are the author’s own and not necessarily the views of the United States government or the Department of the Treasury. This paper has benefitted from conversation with colleagues including Christopher Bradley, Suzanne Garner, and the participants of the 2016 American Society of International Law Biennial Conference of the International Economic Law Interest Group at Georgetown University Law Center, at which an earlier version of this paper was presented. Many thanks are also owed to the excellent editorial staff of the Michigan Journal of International Law.

1. See, e.g., Zeeshan Aleem, *Why Trump’s Plan to Ignore the World Trade Organization Isn’t as Reckless as it Sounds*, VOX (Mar. 8, 2017), <http://www.vox.com/policy-and-politics/2017/3/8/14766228/trump-trade-wto> (“[T]he Trump administration’s skepticism of the WTO isn’t new.”); Catherine A. Rogers, *The Politics of International Investment Arbitration*, 12 SANTA CLARA J. INT’L L. 223, 233 (2013) (noting a pervasive sense of a “legitimacy crisis” in the investment arbitration world); James Surowiecki, *The Financial Page: Trade Agreement Troubles*, THE NEW YORKER, June 22, 2015, at 26 (arguing that investor-state dispute settlement is “outdated and unnecessary” and that “including them in trade agreement undermines the broader case for free trade, by making it look like exactly what people fear—a system designed to put corporate interests above public ones”).

2. These debates—whether we should enshrine rules in international agreements and subject those rules to dispute settlement—are the “substance of government policy.” JAN PAULSSON, DENIAL OF JUSTICE IN INTERNATIONAL LAW 232 (2005). This Article does not wade into them. But we must ensure that the international tools for implementing those policy choices—whatever they may be—produce sound decisions and uphold the legitimacy of the decision-making process.

between China and the Philippines over their territorial seas and a tobacco company's challenge to Uruguay's plain packaging rules for cigarettes—have centered on technical evidence presented by experts.

This is a relatively new feature of international disputes, and we are still in the early days of dealing with its consequences. For the first two centuries of modern international dispute resolution, tribunals and courts sought out expert advice only on rare occasions, usually as a last resort. But reliance on experts has increased dramatically in the last forty years. Human activity itself has become more complex—more scientific, more specialized, more reliant on experts—and so too have disputes that arise out of that activity. At the same time, international agreements have continued evolving to tackle progressively more complex and technical issues, often through express reliance on science, economics, and other specialized fields.

Today, as a result, parties to international disputes routinely put forward expert evidence to support their arguments, and courts and tribunals often appoint their own experts to bridge the gap between law and other fields. Economists and accountants opine on the amount of damages owed to aggrieved investors; scientists evaluate the basis of environmental and health regulations; historians and anthropologists testify on the social, political, and historical contexts of armed conflict and violence; and engineers explain the development of intellectual property and technologies.

Many international tribunals and courts have failed to confront this reality. Most do not ask whether the expert evidence before them even meets a minimum threshold of reliability. And when they do ask, the assessment routinely lacks rigor and transparency. This neglect can lead to reliance on unreliable expert evidence, unfair rejection of evidence that is reliable, or opaque and confused attempts to resolve conflicting expert views. Readers of the resulting decisions are often left wondering about the role, if any, that experts played in resolving the disputes.

The legitimacy of international legal systems—systems predicated on the consent of sovereign nations and entrusted to pass judgment on the public policy choices of those nations—depends on decision making that at least appears analytically sound. International courts and tribunals “must ensure that they are in a position to appreciate the disputed policy choices made by the states appearing before them” through rigorous decision making, defensible analytical frameworks, and thorough appreciation of the facts of each dispute that comes before them.³ This means that the “instruments and processes of international law must provide means for scientific evidence,” and other expert evidence, “to be sifted, understood, and translated into law.”⁴ While decision makers do not “have to under-

3. CAROLINE E. FOSTER, SCIENCE AND THE PRECAUTIONARY PRINCIPLE IN INTERNATIONAL COURTS AND TRIBUNALS: EXPERT EVIDENCE, BURDEN OF PROOF AND FINALITY 77 (2011).

4. Harlan Cohen et al., *An Introduction: Confronting Complexity*, in AM. SOC'Y INT'L L., PROCEEDINGS OF THE 106TH ANNUAL MEETING 1, 1 (2013); see also FOSTER, *supra* note 3, at 77 (“Sound decision-making is essential in an international legal system where submis-

stand a proposition to be justified in believing it,” they do need to “be able to repose a justified trust in the truthfulness and expertise of the person who assures [them] that the proposition is true.”⁵

Importantly, we do find examples where international law has made careful use of expert evidence, and this Article seeks to aggregate best practices from a broad range of substantive areas, including trade and investment, criminal prosecutions, boundary disputes, the law of the sea, post-war and peace commissions, and environmental law. The core of this proposal is that international adjudicators should adopt an analytical framework for determining whether an expert’s testimony is reliable, apply that framework to each expert, clearly set out in the written decision the framework and how it was applied, and take into account the reliability analysis when weighing the evidence presented.

The framework could take the form of a nonexhaustive checklist that would focus on the reliability of an expert’s methodology and the application of that methodology to the facts of a dispute. What is the relevant question and is expert evidence even necessary to answer it? Is the expert’s methodology one that others in the field would use? Has the expert acknowledged and remained within the methodology’s practical limitations? And has the expert presented the methodology and any conclusions in way that is useful to non-experts? Asking those questions and others can help ensure that international law uses expert evidence in a more rigorous, defensible, and transparent manner.

Unlike some other efforts to refine the analytical legitimacy of international adjudication, improving the clarity and transparency with which international arbitrators and judges evaluate expert evidence is in most cases an improvement that can be made by the adjudicators themselves. Existing agreements already vest adjudicators with broad authority and discretion in evaluating expert witness testimony. Harnessing that discretion with a clearly stated decision-making framework has the potential to improve legitimacy and to lead to more correct legal outcomes.

Part I below traces the historical trends in the use of expert evidence in international disputes, from the scattered reliance on experts in the nineteenth and early twentieth centuries to the ubiquity of experts in modern disputes. With that perspective, Part II examines how decision makers have attempted to ensure reliability of the expert evidence that is flooding the evidentiary records of international disputes, while Part III outlines the many problems that still remain. Finally, Part IV proposes a non-exhaustive and nonbinding checklist of questions for analyzing the reliability of any type of expert evidence.

sion to jurisdiction is voluntary and many of the interests at stake are important, sensitive and complex.”).

5. Richard A. Posner, *The Law and Economics of the Economic Expert Witness*, 13 J. ECON. PERSP. 91, 96 (1999).

I. HISTORICAL TRENDS IN THE USE OF EXPERT EVIDENCE

Exploring the historical context is critical to understanding how international courts and tribunals can best make use of expert evidence. Although commonplace today, reliance on expert evidence in international disputes is a relatively new phenomenon, rare until the last forty years.

The first hundred years of the modern era of international dispute settlement—between the 1794 Jay Treaty⁶ and World War I—provide few salient examples of reliance on expert testimony. Although more than 300 tribunals operated during this time,⁷ and the resulting decisions have been described as producing a “vast international arbitral jurisprudence,”⁸ the century provides few salient examples of reliance on expert testimony.

The peace treaties and other international instruments signed in the wake of the World Wars led to a modest increase in reliance on expert testimony.⁹ The various Mixed Arbitral Tribunals established under the World War I peace treaties used expert witnesses particularly frequently.¹⁰ In *Caisse d'Assurances des Glaceries v. Germany*, for example, the Belgian-German tribunal sought expert testimony on the value of a factory

6. Arbitral or judicial settlement of international disputes has been described as “irregular and spasmodic” before the 1794 signing of the Jay Treaty between Great Britain and the United States. See JACKSON RALSTON, INTERNATIONAL ARBITRATION FROM ATHENS TO LOCARNO 191 (1929). Because that settlement “assumed a certain regularity and system” after the Jay Treaty, the Treaty is generally considered the beginning of the modern era of international dispute settlement. *Id.* at vii, 191.

7. See RALSTON, *supra* note 6, at 345–54 (listing 304 “arbitral and other judicial tribunals functioning between nations” from 1794 to before the World War I tribunals that began in 1919); see also EVANS DARBY, INTERNATIONAL TRIBUNALS: A COLLECTION OF THE VARIOUS SCHEMES WHICH HAVE BEEN PROPOUNDED, AND OF INSTANCES SINCE 1815 285–304 (1899) (listing 158 international tribunals between 1815 and 1899); Paulsson, *supra* note 2, at 131 n.1 (noting that some of these tribunals resolved multiple disputes).

8. PAULSSON, *supra* note 2, at 131.

9. See GILLIAN M. WHITE, THE USE OF EXPERTS BY INTERNATIONAL TRIBUNALS 80–81 (1965) (noting that treaties and rules of procedure increasingly “[ma]de use of expert help in the form of testimony, reports and investigations on the spot” and that “several of the tribunals established after the world wars . . . have availed themselves of such powers”). Several World War I peace treaties provided for the use of experts. See *id.* at 50 (citing as examples Article 56 of the Rules of the Belgian-German tribunal, Articles 57 and 60 of the Rules of the Franco-German tribunal, Article 3 of the Agreement of June 30, 1921 Between the United States of America and Norway for the Submission to Arbitration of Certain Claims of Norwegian Subjects, and Article 3 of the Arbitration Protocol between France and Haiti of September 10, 1913).

10. See DURWARD V. SANDIFER, EVIDENCE BEFORE INTERNATIONAL TRIBUNALS 325 n.137 (1975) (citing *Huret v. État Allemand* (Franco-German Tribunal), Vol. 1, RECUEIL DES DÉCISIONS DES TRIBUNAUX ARBITRAUX MIXTES 98, 488 (1921); *Lirens v. État Allemand* (German-Belgian Tribunal), 2 RECUEIL DES DÉCISIONS DES TRIBUNAUX ARBITRAUX MIXTES 82; *Companie d'Electricité de Sofia et de Bulgarie v. État Bulgare et Municipalite de Sofia* (Bulgarian-Belgian Tribunal), 3 RECUEIL DES DÉCISIONS DES TRIBUNAUX ARBITRAUX MIXTES 308, 323 (1923)). Many of the experts involved in these Tribunals were engaged to assist in valuations. For example, the Franco-German tribunal in *Huret v. State of Germany* appointed an expert to value a specific make and model of a car that had been requisitioned. See 1 RECUEIL DES DÉCISIONS DES TRIBUNAUX ARBITRAUX MIXTES 98, 488; see also WHITE, *supra* note 9, at 51–52.

expropriated by the Germans, including the “sum required . . . to construct [it] in Germany,” what its “current value” would have been, and “what profits it might have earned to date.”¹¹

Although the commissions established under the 1947 peace treaties,¹² the 1951 peace treaties with Japan,¹³ and the early 1950s Bonn-Paris Conventions with Germany¹⁴ were broadly empowered to consider expert evidence or to develop their own procedures,¹⁵ it does not appear that the commissions established after World War II made widespread use of experts.¹⁶ As an example, the Arbitral Commission on Property, Rights, and Interest in Germany, established in 1954 and described as “one of the most active of the . . . international courts” of the mid-century period, adjudicated more than one hundred cases but referred a matter to experts in only one reported decision.¹⁷ In that case, *Greece and the Firm Apostolidis v. Germany*, experts determined the going price of chrome in Germany to form the basis of the Commission’s compensation calculation.¹⁸ Of the World War II commissions, the Franco-Italian Commission is the

11. 3 RECUEIL DES DÉCISIONS DES TRIBUNAUX ARBITRAUX MIXTES, *supra* note 10, at 261, 270; *see also* WHITE, *supra* note 9, at 51-52.

12. *See, e.g.*, Treaty of Peace with Italy art. 83(3), Feb. 10, 1947, 61 Stat. 1245 (empowering the Italian Commissions to “adopt [] rules conforming to justice and equity”); Rules of Procedure of the Franco-Italian Commission art. 14, *translated in* WHITE, *supra* note 9, at 53 (“The Commission may decide . . . to arrange for the carrying out of any expert enquiry with the co-operation of all necessary technicians, interpreters or translators.”).

13. *See, e.g.*, Rules of Procedure of the United States-Japanese Property Commission art. 13(E), *in* RICHARD B. LILICH & GORDON A. CHRISTENSON, INTERNATIONAL CLAIMS: THEIR PREPARATION AND PRESENTATION app. F, at 144 (1962) (“The [United States-Japanese Property] Commission may appoint an expert and request him to submit an opinion in writing.”).

14. Convention on the Settlement of Matters Arising out of the War and the Occupation, Oct. 23, 1954, [1955] 6 U.S.T. & O.I.A. 4411, T.I.A.S. No. 3425 (“The Commission shall have the power . . . to request expert opinion.”); Rules of Procedure of the Arbitral Commission on Property, Rights, and Interest in Germany, Rule 22(a), *in* 1 DECISIONS OF THE ARBITRAL COMMISSION ON PROPERTY, RIGHTS, AND INTEREST IN GERMANY 196 (1958) [hereinafter Rules of Procedure of the German Arbitral Commission] (“The Commission shall have power, upon application of a party or of its own motion . . . to request expert opinion.”).

15. *See* WHITE, *supra* note 9, at 52-54. Although all the 1947 peace treaties “included identical provisions for the settlement of disputes . . . only those in the peace treaty with Italy were put into effect.” *Id.*

16. *See* WHITE, *supra* note 9, at 55 (noting that “none of the published decisions” of the United States-Japanese Property Commission “mention that the Commission ever made use of the power”). This may have been because “in no case did it prove necessary to consider detailed evidence of damages.” Lionel M. Summers & Arnold Fraleigh, *The United States-Japan Property Commission*, 56 AM. J. INT’L L. 407, 411 (1962).

17. WHITE, *supra* note 9, at 57 (citing *Government of the Kingdom of Greece and the firm Apostolidis v. Federal Republic of Germany*, 3 DECISIONS OF THE ARBITRAL COMMISSION ON PROPERTY, RIGHTS, AND INTERESTS IN GERMANY 329, 364 (1960)).

18. *Government of the Kingdom of Greece and the firm Apostolidis v. Federal Republic of Germany*, 3 DECISIONS OF THE ARBITRAL COMMISSION ON PROPERTY, RIGHTS, AND INTERESTS IN GERMANY 329, 364 (1960).

only one cited as providing “many examples of the use of independent experts.”¹⁹

The standing dispute resolution institutions of the era also rarely relied on experts. The Permanent Court of Arbitration did not rely on expert testimony for its first several decades, even though it possessed express authority to appoint its own experts.²⁰ In its eighteen years of existence, between 1922 and 1940, the League of Nations’s Permanent Court of International Justice (PCIJ) sought out expert testimony in just one dispute, the widely cited *Factory at Chorzów* case.²¹ And in the post-war era, the International Court of Justice (ICJ), the heir to the PCIJ, continued the trend, relying on expert testimony only once, in the 1949 *Corfu Channel* case.²²

There were exceptions. One class of cases in particular—boundary disputes—consistently bucked the general trend.²³ From the early days of modern international dispute resolution, decision makers regularly relied on expert evidence on geography and hydrogeology,²⁴ including the St. Croix River dispute between Great Britain and the United States in 1798.²⁵ Other than questions of delineating sovereign boundaries, there

19. WHITE, *supra* note 9, at 53. A prominent example from the Commission is the Ousset Claim. See 5 RECUEIL DES DÉCISIONS DE LA COMMISSION DE CONCILIATION FRANCO-ITALIENNE 36; 22 I.L.R. 312 (1955). During the first decade of existence, neither the European Commission of Human Rights nor the European Court of Human Rights exercised its authority to rely on expert testimony. See WHITE, *supra* note 9, at 62, 66.

20. See WHITE, *supra* note 9, at 36, 74, 138. In one case, *Lighthouses Arbitration (Fr. v. Greece)*, the Permanent Court of Arbitration ordered expert testimony, but the parties settled the claims before experts were appointed. 12 R.I.A.A. 155; 23 I.L.R. 659 (1956).

21. *Factory at Chorzów (Ger. v. Pol.) (Indemnity)*, 1928 P.C.I.J. (ser. A) No. 17, at 136-51); see also WHITE, *supra* note 9, at 40.

22. See *Corfu Channel (U.K. v. Alb.)*, Merits, 1949 I.C.J. Rep. 4 (Apr. 9) (relying on tribunal-appointed experts to determine liability); *Corfu Channel (U.K. v. Alb.)*, Judgment, 1949 I.C.J. Rep. 244 (Dec. 15) (relying on tribunal-appointed experts to determine the amount of compensation owed).

23. Adjudicators under Roman law relied on experts in land disputes, see ARTHUR ENGLEMAN, A HISTORY OF CONTINENTAL CIVIL PROCEDURES 361-62 (1928), as did Italian and French courts since the middle ages, with codification of expert procedures in a 1667 French ordinance, see WHITE, *supra* note 9, at 16.

24. See, e.g., A. de LAPRADELLE & N. POLITIS, 2 RECUEIL DES ARBITRAGES INTERNATIONAUX 574, 579 (2d ed. 1957) (*Eastern Boundary (S. Afr. v. Orange Free State)* (1870)); *Manica Plateau Arbitration (Gr. Brit. v. Port.)*, 28 R.I.A.A. 283, 308 (1897)); *Cordillera of the Andes Boundary Case (Arg. v. Chile)*, 9 R.I.A.A. 29, 39-40 (Perm. Ct. Arb. 1902); *L’Oeil de la Mer (Austria v. Hung.)* (1902), 8 REV. DE DROIT INT’L ET DE LÉGISLATION COMP. 162 (2d ser. 1906); 3 MARTENS NOUVEAU RECUEIL GÉNÉRALE DE TRAITÉS (ser. 71) 465; *Costa Rica-Panama Boundary (Costa Rica v. Pan.)*, 11 R.I.A.A. 518, 539-40 (1914); *Honduras Borders (Guat. v. Hond.)*, 2 R.I.A.A. 1307, 1353 (Int’l Central American Trib., 1933); *Gran Chaco Dispute (Bol. v. Para.)*, 3 R.I.A.A. 1817, 1822-25 (1938).

25. See JOHN BASSETT MOORE, HISTORY AND DIGEST OF THE INTERNATIONAL ARBITRATIONS TO WHICH THE UNITED STATES HAS BEEN A PARTY 23-24, 29-30 (1898). The expert testimony—that of land surveyors and astronomers—may have played a significant role in the outcome, as the proceedings were twice postponed while their surveys and reports were completed. *Id.* at 23-24. The award was rendered only ten days after receipt of their final map. *Id.* at 24. The map was annexed to the award. *Id.* at 30.

were relatively few examples of reliance on expert evidence—some related to valuations, science, document authentication, and engineering, for example—in the nineteenth century and the twentieth century through the 1970s.²⁶

The trend in the last forty years has been very different, as tribunals and courts have dramatically increased the frequency and significance of their reliance on expert evidence. Many factors appear to have contributed to the trend, which parallels the increased reliance on experts in other decision-making processes, including international governance outside legal dispute resolution.²⁷ The first order explanation is simply that human activity itself has become more complex—more scientific, more specialized, more reliant on experts—and so too have disputes that arise out of that activity.²⁸ But international law itself has played a role, too.

The more robust procedural and evidentiary frameworks in international agreements may also explain some of the rise in reliance on experts. Express authorization to rely on expert testimony gradually became a ubiquitous feature of procedural rules of standing tribunals and courts in the mid-twentieth century.²⁹ International dispute settlement is also in-

26. See, e.g., *Lehigh Valley Railroad Co. (U.S. v. Ger.)*, 8 R.I.A.A. 104, 121 (1932) (analyzing over one thousand pages of evidence submitted by document authentication experts); *I.R. Clark (U.S. v. Mex.)*, 4 R.I.A.A. 415, 415 (1928) (noting that the Mexico-U.S. Commission would give “careful consideration” to the “expert testimony” of a physician); *Marguerite de Joly de Sabla (U.S. v. Pan.)*, 6 R.I.A.A. 358, 362, 367 (1933) (relying in part on the testimony of a licensed surveyor); *Fur Seal Arbitration (U.S. v. Gr. Brit. 1892)*, MOORE, *supra* note 25, at 755, 813–14 (examining the testimony of fisheries experts).

27. See Monika Ambrus et al., *The Role of Experts in International and European Decision-Making Processes: Setting the Scene*, in *THE ROLE OF “EXPERTS” IN INTERNATIONAL AND EUROPEAN DECISION-MAKING PROCESSES: ADVISORS, DECISION MAKERS OR IRRELEVANT ACTORS?* 1, 5–6 (M. Ambrus et al. eds., 2014) (“Experts are called upon to provide input into ever more complex decision-making processes. They are involved in policy making, law-making, and implementation and enforcement efforts.”).

28. See Anna Riddell, *Scientific Evidence in the International Court of Justice: Problems and Possibilities*, in 20 *FINNISH J. OF INT’L L.* 229, 230 (2009) (“Major technological and scientific developments in recent decades have ensured that scientific knowledge has developed an important role in the administration of law.”); Dushyant Dave, *Contemporary Practice in the Conduct of Proceedings: Techniques for Eliciting Expert Testimony – How Party-Appointed Experts Can Be Made Most Useful*, in *INTERNATIONAL ARBITRATION 2006: BACK TO BASICS?* 813, 818 (Albert Jan van den Berg ed., 2006) (“[I]nternational tribunals have become much more dependent on expert testimony, and the accelerating process of scientific and technological advance enhances this tendency.”).

29. See, e.g., Consolidated Version of the Protocol on the Statute of the Court of Justice of the Coal and Steel Community, art. 25, Apr. 18, 1951 (“The Court may at any time entrust any individual, body, authority, committee or other organisation it chooses with the task of holding an inquiry or giving an expert opinion; to this end it may compile a list of individuals or bodies approved as experts.”); Maarten Bos, *The Franco-Italian Conciliation Commission*, 22 *NORDIC J. OF INT’L L.* 133, 154 (1952)

Article 14, paragraph 3, [of the Rules of Procedure of the Franco-Italian Conciliation Commission] confers upon the Commission the right to have examined by experts nominated by itself the reports of the experts nominated by the parties.

creasingly open to nongovernmental groups, international organizations, academics, and the public at large, including through the submission of non-disputing party submissions and amicus curiae briefs, many of which may present expert-like opinions.³⁰

At the same time, this is partly because international agreements over the years have evolved to tackle ever more complex issues, themselves becoming more technical, more specialized, and more tailored to the problems that they seek to address.³¹ Countries negotiating trade agreements, once concerned primarily with reducing tariff barriers and quotas, have shifted focus to other internal laws and regulations, such as those concerning health, safety, and the environment, that may act as unfairly trade restricting or distorting.³² Agreements also often require or en-

Under paragraph 4 of this Article the Commission may visit the places concerned in a dispute and may hear its own experts there. The Agents or their substitutes are, by paragraph 5, entitled to be present.

30. See Joost Pauwelyn, *The Use of Experts in WTO Dispute Settlement*, 51 INT'L COMP. L.Q. 325, 325 (2002) (“[A]n increasing number of ‘outsiders’ or amici curiae, such as NGOs, but also industry and academics, have pressed their (expert) opinion on WTO panels and the Appellate Body.”); OFF. OF THE U. S. TRADE REPRESENTATIVE: THE TRANS-PAC. PARTNERSHIP art. 9.23(3), <https://ustr.gov/sites/default/files/TPP-Final-Text-Investment.pdf> (last visited Feb. 5, 2017) (“After consultation with the disputing parties, the tribunal may accept and consider written amicus curiae submissions regarding a matter of fact or law within the scope of the dispute that may assist the tribunal in evaluating the submissions and arguments of the disputing parties from a person or entity that is not a disputing party but has a significant interest in the arbitral proceedings.”); NAFTA Free Trade Comm’n, *Statement of the Free Trade Commission on Non-Disputing Party Participation*, (Oct. 7, 2004) <http://naftaclaims.com/commissionfiles/Nondisputing-en.pdf> (recommending that NAFTA investor-state tribunals adopt certain procedures regarding non-disputing party submissions). A number of investment arbitration decisions have relied on amicus curiae submissions presenting scientific and other technical evidence. See, e.g., Philip Morris Brands Sarl v. Uruguay, ICSID Case No. ARB/10/7, Award ¶¶ 394, 396 (2015) (relying on a World Health Organization submission); Methanex Corp. v. United States, Decision of the Tribunal on Petitions from Third Persons to Intervene as “Amici Curiae” ¶ 53 (Jan. 15, 2001) (determining that a tribunal has the authority to accept amicus written submissions under Article 15(1) of the UNICTRAL Arbitration Rules).

31. See Kate Miles, *Climate Change: Trading, Investing and the Interaction of Law, Science and Technology*, in SCIENCE AND TECHNOLOGY IN INTERNATIONAL ECONOMIC LAW: BALANCING COMPETING INTERESTS 155, 155 (B. Mercurio & Kuei-Jung Ni eds., 2014) (“The ever-increasing complexity of the issues with which international law is required to grapple has recently seen, among other effects, a notable visibility of science and technology in international law.”).

32. See Lukasz Gruszczynski, *Science and the Settlement of Trade Disputes in the World Trade Organization*, in SCIENCE AND TECHNOLOGY IN INTERNATIONAL ECONOMIC LAW: BALANCING COMPETING INTERESTS 11, 24 (B. Mercurio & Kuei-Jung Ni eds., 2014) (“GATT 1947 was predominantly concerned with tariff barriers and quotas. Its success in reducing these types of obstacle has shifted the attention of countries to market access problems caused by internal measures (i.e., non-tariff barriers). In consequence, potential disputes under WTO law have become not only more complex (than tariff barriers and quotas), but also more politically contentious.”); Robert Howse, *Democracy, Science, and Free Trade: Risk Regulation on Trial at the World Trade Organization*, 98 MICH. L. REV. 2329, 2329 (2000) (“Traditionally, free trade rules were about constraining border measures such as tariffs and quantitative restrictions on imports. Increasingly, however, such rules include requirements and constraints addressed directly to domestic regulation.”).

courage countries to adopt international standards or establish scientific evidence requirements for certain types of technical (especially environmental, health, and safety) regulation.³³ International legal systems that have emerged in the last forty years, some with inherently technical mandates and legal norms that are inextricably bound up with questions of specialized knowledge, routinely elicit and rely on expert testimony.³⁴ Recent agreements also contain procedural provisions that contemplate tribunal-appointed experts on specific technical issues³⁵ or that otherwise encourage reliance on expert evidence.³⁶ Partly as a result of these provisions, international disputes increasingly arise not only with respect to past injury but also risks of future harm—to human health or the environment, for example—and expert testimony plays a significant role in those disputes.³⁷

33. See Wouter G. Werner, *The Politics of Expertise: Applying Paradoxes of Scientific Expertise to International Law*, in *THE ROLE OF “EXPERTS” IN INTERNATIONAL AND EUROPEAN DECISION-MAKING PROCESSES: ADVISORS, DECISION MAKERS OR IRRELEVANT ACTORS?* 44, 55 (M. Ambrus et al. eds., 2014) (“Legal experts themselves . . . are also increasingly dependent on experts from other disciplinary fields. The fragmentation of international law has created functional subfields where legal norms can often only be applied on the basis of scientific expert knowledge.”); FOSTER, *supra* note 3, at 114 (“A significant development in the taking of scientific evidence by international courts and tribunals is the system that has developed in the WTO. The WTO system for taking expert evidence was devised as a response to the needs of the dispute resolution process in cases involving complex scientific questions.”).

34. See, e.g., Mark F. Rosenberg & Michael A. Cheah, *Arbitrating Environmental Disputes*, 16 *ICSID REV.* 39, 42 (2001) (“[T]here is a body of treaties that deal specifically with environmental concerns and may provide for arbitration in the event of disagreement between the signatories.”). Several disputes arising under the 1982 United Nations Convention on the Law of the Sea Treaty, for example, have relied heavily on expert evidence. See, e.g., *Phil. v. China*, Award, PCA Case Repository 2013-19 (2016) (discussing the terms and necessary skillset of a potential appointed expert); *Bay of Bengal Maritime Boundary Arb. (Bangl. v. India)*, Award, PCA Case Repository 2010-16 (2014) (appointment of an expert hydrographer); *Guy. v. Surin.*, Award, PCA Case Repository 2004-04 (2007) (detailing the technical responsibilities of the appointed expert).

35. See, e.g., OFF. OF THE U. S. TRADE REPRESENTATIVE: THE TRANS-PAC. PARTNERSHIP art. 28.1.5, <https://ustr.gov/sites/default/files/TPP-Final-Text-Dispute-Settlement.pdf> (last visited Feb. 5, 2017) (“Role of Experts: At the request of a disputing Party, or on its own initiative, a panel may seek information and technical advice from any person or body that it deems appropriate, provided that the disputing Parties agree and subject to any terms and conditions agreed by the disputing Parties. The disputing Parties shall have an opportunity to comment on any information or advice obtained under this Article.”); Agreement Between Canada and the Republic of Peru for the Promotion and Protection of Investments art. 42; *Can.-Peru*, entry into force June 20, 2007, 2007 Can. T.S. No 10 (2006); Treaty Between the United States and the Oriental Republic of Uruguay Concerning the Encouragement and Reciprocal Protection of Investment art. 32, U.S.-Uru., (Nov. 4, 2005), S. Treaty Doc. No. 109-9; United States-Singapore Free Trade Agreement art. 15.23, U.S.-Sing., May 6, 2003, P.L. 108-78.

36. See, e.g., Agreement on the Application of Sanitary and Phytosanitary Measures, 33 I.L.M. 15 (1994), art. 11.2 [hereinafter SPS Agreement] (“In a dispute under this Agreement involving scientific or technical issues, a panel should seek advice from experts chosen by the panel.”).

37. See FOSTER, *supra* note 3, at xiii, xiv, xvii.

Together, these developments have significantly increased the role that experts play in resolving international disputes, in a broad range of disputes and a wide variety of forums. Where disputes were once focused on primarily legal issues with relatively simple fact patterns—the propriety of war-time actions and expropriations of relatively modest industrial facilities³⁸—disputes today often involve complex analyses of health, safety, or environmental risks, or of the economic fates of enormous businesses. Economists and accountants opine on the amount of damages owed,³⁹ scientists play prominent roles in evaluating the basis of environmental and health regulations,⁴⁰ and engineers opine on the operation of industrial facilities and extractive concessions.⁴¹ In international criminal cases, historians and anthropologists testify on the social, political, and historical contexts of armed conflict and violence.⁴² And many disputes require the piecing together of testimony from multiple experts. An investment dispute involving an industrial project might require an engineer to evaluate the project’s output capacity, an economist to evaluate income streams, and an accountant and tax specialist to assess the associated regulatory and business costs.⁴³

The complexity of international disputes means that when decision makers rely on expert evidence, they rely heavily on it.⁴⁴ Expert evidence

38. See, e.g., notes 15-20 and accompanying text.

39. See, e.g., *Starrett Housing Corp. v. Iran*, Interlocutory Award No. ITL 32-24-1, 4 Iran-U.S. Cl. Trib. Rep. 122, 157 (1983).

40. See, e.g., notes 61-66 and accompanying text.

41. See, e.g., *Harza v. Iran*, Interlocutory Award No. ITL 14-97-2, 2 Iran-U.S. Cl. Trib. Rep. 68, 71 (1983) (appointing experts in geology, hydraulics, and business practices and administration of consulting engineering contracts).

42. See Doris Buss, *Expert Witnesses and International War Crimes Trials: Making Sense of Large-Scale Violence in Rwanda*, in *NARRATIVES OF JUSTICE IN AND OUT OF THE COURTROOM* 26 (Dubravka Zarkov & Marlies Glasius eds., 2014); see also Sean D. Murphy, *The Experience of the Eritrea-Ethiopia Claims Commission* 106 *PROCEEDINGS OF THE ANNUAL MEETING (AM. SOC’Y INT’L L.)* 237, 240 (2012) (noting that the Eritrea-Ethiopia Claims Commission relied on expert witnesses including “psychiatrists, medical doctors, retired military officers from third states, and various military and explosives experts” (citations omitted)); Avi Singh, *Expert Evidence*, in *PRINCIPLES OF EVIDENCE IN INTERNATIONAL CRIMINAL JUSTICE* 600 (K. Khan et al. eds., 2010) (“As in domestic criminal proceedings, expert witnesses are critical, seemingly more than eye witnesses, to adduce evidence in inherently complex international trials.”).

43. See Sergey Ripinsky & Kevin Williams, *Damages in International Investment Law* 175 (2008) (noting that it may be “sometimes essential” to appoint experts with backgrounds in accounting, finance, and “the industry or market in which the company operates” in valuing businesses).

44. See FOSTER, *supra* note 3, at 77 (“The complexity of the science requires heavy consultation with experts, and considerable reliance on their testimony.”); Thomas Cottier, *Risk Management Experience in WTO Dispute Settlement*, in *GLOBALIZATION AND THE ENVIRONMENT: RISK ASSESSMENT AND THE WTO* 41, 51 (D. Robertson & A. Kello eds., 2001) (arguing that, in *EC—Hormones*, *Australia—Salmon*, and *Japan Agricultural Products II*, “panels strongly relied upon expert views, and could not have rendered their assessment and recommendation without such independent advice of individual experts”); RIPINSKY & WILLIAMS, *supra* note 43, at 174 (characterizing the quantification of damages as a “very complex

often has wide-ranging and significant influence on decision makers and may even determine the outcome of key issues in a dispute.⁴⁵ That is true when the governing legal instruments refer to inherently technical concepts, such as the World Trade Organization's (WTO) Agreement on the Application of Sanitary and Phytosanitary Measures (SPS Agreement) and its references to risk assessments, but also when the legal instruments do not.⁴⁶ Expert evidence in international disputes may inform the meaning of terms in international agreements—where the legal text refers to the reasonableness or necessity of certain actions, for example—and shape the way that courts and tribunals view the policy choices of states in every area of governance.⁴⁷ Even an expert's "back of the envelope" conclusions

exercise requiring special knowledge," making "the involvement of valuation experts in arbitral proceedings practically inevitable" and giving those experts a "singularly important role . . . in the matter of assessing compensation"); Dana Freyer, *Assessing Expert Evidence, in* *Leading Arbitrators' Guide to International Arbitration* 429, 440 (2008) ("[T]he expert with technical expertise . . . may effectively act as the final authority on matters of great importance to the case."); Miles, *supra* note 31, at 155-56 ("It is the sum of these incidences [that have led to the increasing complexity of international disputes] . . . that has wider implications . . . for . . . the proposition that science and technology are influencing the development of international law in fundamental ways, that they play an increasingly pivotal role in trade and investment regulation and dispute settlement and, more specifically, that they will also be central to the treatment of climate-related issues within international economic law."); Dave, *supra* note 28, at 817 ("[A]rbitrators and even lawyers have had to become proficient in weighing competing expert evidence.").

45. See Miles, *supra* note 31, at 160 ("In the context of international litigation, it is the use of scientific expert evidence to support the legal arguments presented that can be particularly influential on decision-makers and, indeed, determine the outcome of a dispute."); *id.* ("Inevitably, . . . in health and or environment-related cases, there is a heavy reliance on the input of experts as witnesses or as advocates themselves."); *id.* at 161 ("[S]cience-based evidence, its presentation[,] and the authority of the individual experts will be central to a court's decision."); Catherine E. Foster, *New Clothes for the Emperor? Consultation of Experts by the International Court of Justice*, 5 J. INT'L DISPUTE SETTLEMENT 139, 140 (2014) ("[E]xperts wield potentially significant influence in the determination of the scope of sovereign States' future freedoms."); MICHELLE T. GRANDO, EVIDENCE, PROOF, AND FACT-FINDING IN WTO DISPUTE SETTLEMENT 340 (2009) ("[WTO] panels . . . usually accord [expert] . . . opinions great weight."); *id.* (arguing that the panel in *EC — Hormones* "relied entirely on the opinion of the experts to support its finding" that naturally occurring hormones are comparable to hormones administered for therapeutic or zootechnical purposes and therefore subject to the disciplines of Article 5.5 of the SPS Agreement); *id.* at 341 (concluding that the *Japan-Apples* panel "relied entirely upon the opinion of the experts in addressing" whether a class of apples could act as a disease-transmission vector); see also United States—Continued Suspension of Obligations in the EC-Hormones Dispute, Appellate Body Report, WT/DS320/AB/R (2008), ¶ 480 ("Experts appointed by a panel can significantly influence the decision-making process.").

46. See GRANDO, *supra* note 45, at 340 ("In most [WTO] cases where experts were consulted, particularly those brought under the SPS Agreement, the experts' opinions had a clear impact on the findings of the panels.").

47. See Miles, *supra* note 31, at 77-78

The insights offered by a scientific expert will help determine the application of a legal concept such as 'necessity' or 'reasonableness' in the case at hand – and in the course of this process the scientific expert's advice will also come to shape the development of the conventions and usages on which the established meaning of

made on the witness stand may carry significant weight with a court or tribunal.⁴⁸ Expert reports, exhibits, and testimony contribute significantly to the tens of thousands of pages that routinely pile up in international disputes.⁴⁹ This trend shows no signs of stopping.⁵⁰

The use of experts in international law has been especially prevalent in trade and investment disputes. In the more than fifty years of the multilateral trading system under the 1947 General Agreement on Trade and Tariffs (GATT), only one panel requested an expert opinion⁵¹ out of approximately 300 disputes,⁵² and in that instance did not rely on it.⁵³ Early cases in the WTO also tended to avoid reliance on expert evidence.⁵⁴ Yet, today reliance on expert evidence has become a standard practice in the

the legal concept of ‘necessity’ or ‘reasonableness’ in the context of the provision in question will be based.

48. See FOSTER, *supra* note 3, at 120; see also Joost Pauwelyn, *Expert Advice in WTO Dispute Settlement*, in TRADE AND HUMAN HEALTH AND SAFETY 235, 248 (George A. Bermann & Petros C. Mavroidis eds., 2006).

49. See Lucy Reed, *The Kaplan Lecture 2012: Arbitral Decision-Making: Art, Science or Sport?* 12 (2012), http://arbitration-icca.org/media/1/13581569903770/reed_tribunal_decision-making.pdf (“With flexible rules and no page limits in international arbitration . . . we see records with 500-page memorials, scores of witness statements and expert reports, and thousands of exhibits. How can we expect the arbitrators to find a trail up such a mountain of a record?”); see also CHESTER BROWN, A COMMON LAW OF INTERNATIONAL ADJUDICATION 112 (2007) (noting that parties to ICJ disputes submit “large volumes of complex technical, scientific or financial material”); FOSTER, *supra* note 3, at 108 (noting “the mass of scientific evidence” in the ICJ’s *Gabčikovo-Nagymaros* case); Christophe Bonneuil & Les Levidow, *How Does the World Trade Organization Know? The Mobilization and Staging of Scientific Expertise in the GMO Trade Dispute*, 42 SOC. STUD. SCI. 75, 77 (2012) (citing the more than 1,000 pages of “discussions” regarding expert evidence in Panel Report, *EC — Approval and Marketing of Biotech Products*, WTO/DS291 (2006)).

50. See FOSTER, *supra* note 3, at 132 (“Expert evidence is likely to be given a higher profile in international judgments and awards in future.”); Foster, *supra* note 45, at 139 (“Disagreements over science feature with growing regularity in international disputes.”).

51. See Panel Report, *Thailand — Cigarettes*, ¶ 27, WTO Doc. DS10/R - 37S/200 (Oct. 5, 1990) (requesting advice from the World Health Organization); see also Pauwelyn, *supra* note 48 at 235 (George A. Bermann & Petros C. Mavroidis eds., 2006) (noting that there has only been “one single resort to experts by GATT 1947 panels”).

52. *WTO in Brief: Part 3 – The WTO Agreements*, WORLD TRADE ORGANIZATION, 5 (2014), https://wto.org/english/res_e/doload_e/inbr_e.pdf.

53. See Gruszczynski, *supra* note 32, at 16–17.

When deciding the [*Thailand – Cigarettes*] case, the GATT panel took for granted that the Thai measure was backed up by sufficient scientific evidence. At the same time, it did not address the claims of Thailand that US cigarettes were actually more harmful than Thai ones (owing to the presence of additives and the specific process used in their production).

54. *Id.* at 17-18 (citing *US-Gasoline* and *US- Shrimp* as examples of the WTO Appellate Body and panels avoiding reliance on expert scientific evidence).

WTO,⁵⁵ which since 1994 has routinely relied on such evidence.⁵⁶ And WTO observers have called for more frequent use of experts in particular areas, including economics.⁵⁷

After its initial reticence, the WTO has consistently relied on scientific experts, in part because a number of WTO provisions expressly require consistency with scientific norms.⁵⁸ Panels have in particular relied heavily on scientific expertise in claims submitted under the SPS Agreement⁵⁹ and the Agreement on Technical Barriers to Trade, and when analyzing a number of exceptions to WTO obligations, such as those in the General Agreement on Tariffs and Trade, and the General Agreement on Trade in Services.⁶⁰ No international legal text goes further in expressly relying on technical evidence than the SPS Agreement, which relies in significant part on science to distinguish between product health measures that are consistent with the Agreement and those that are not.⁶¹ The Agreement

55. See Jessica Lawrence, *The Structural Logic of Expert Participation in WTO Decision-Making Processes*, in *THE ROLE OF "EXPERTS" IN INTERNATIONAL AND EUROPEAN DECISION-MAKING PROCESSES: ADVISORS, DECISION MAKERS OR IRRELEVANT ACTORS?* 173, 176 (M. Ambrus et al. eds., 2014) ("Reliance on scientific expertise has become a standard practice at the WTO."); FOSTER, *supra* note 3, at 133 (citing *Japan – Measures Affecting the Importation of Apples and United States – Import Prohibition of Certain Shrimp and Shrimp Products* as two examples of reliance by WTO panels on expert evidence); Pauwelyn, *supra* note 48, at 235 (explaining that, in the first six years of the WTO's existence, "six panels appointed scientific experts," "[t]wo panels requested expert advice from other international organisations," and "one panel appointed a linguistic expert" and that "very often parties to a WTO dispute also nominate experts on their delegation, be they lawyers, economists, scientists or linguists" and NGOs, industry, and academics "have pressed their (expert) opinion on WTO panels").

56. See Pauwelyn, *supra* note 30, at 325 (noting 12 cases between 1994 and 2002).

57. See, e.g., GRANDO, *supra* note 45, at 342 ("The significant role which experts have played in scientific disputes suggests that more consideration should be given to the idea of appointing economic experts when panels are asked to make complex economic determinations that fall outside their area of expertise."); Andre Sapir & Joel Trachtman, *Subsidization, Price Suppression, and Expertise: Causation and Precision in Upland Cotton*, 7 *WORLD TRADE REV.* 183, 205–07 (2008) (arguing that panels should use economic experts in cases requiring complex economic determinations).

58. See Lawrence, *supra* note 55, at 176 ("[C]onformity with scientific norms is an explicit criterion of compliance with a number of WTO obligations."); Gruszczynski, *supra* note 32, at 23 ("[S]cience is an important element in all health and environment-related trade disputes, irrespective of which agreement they are brought under."); see also Lawrence, *supra* note 55, at 174 ("[e]xperts are not something that can be 'added to' or 'subtracted from' the WTO in order to tweak input or output legitimacy. Instead their [inputs] . . . are the fabric from which the organization is cut."); GRANDO, *supra* note 45, at 338 (2009) ("In [WTO] cases where technical matters are at issue, experts may play an important role in ensuring an accurate decision.").

59. See Gruszczynski, *supra* note 32, at 22 ("[S]cience has become a central benchmark for the assessment of SPS measures.").

60. See Pauwelyn, *supra* note 48, at 235.

61. See Vern R. Walker, *Keeping the WTO from Becoming the World Trans-Science Organization: Scientific Uncertainty, Science Policy, and Factfinding in the Growth Hormones Dispute*, 31 *CORNELL INT'L L.J.* 251, 253 (1998) ("The central strategy of the SPS Agreement is to use science to distinguish between those sanitary measures consistent with the Agreement and those in violation of the Agreement."); see also Panel Report, *EC — Approval and*

helps to address the possibility that a WTO member might use health concerns as a pretense to impose restrictions on imported goods in order to protect its own industry from foreign competition.⁶² Article 5, for example, requires governments to ensure that certain health and safety requirements are “based on an assessment, as appropriate to the circumstances, of the risks to human, animal or plant life or health, taking into account risk assessment techniques developed by the relevant international organizations.”⁶³

The increasing number of investor-state arbitrations under bilateral investment treaties and free trade agreements in the last twenty years has contributed significantly to the overall trend in expert evidence. This significant reliance is nowhere more apparent than in valuation determinations in investment arbitrations.⁶⁴ But experts routinely play prominent roles in evaluating the scientific basis of health and environmental regulation in investment disputes, too. Challenged regulations have included fuel additive bans,⁶⁵ plain packaging requirements for cigarettes,⁶⁶ public announcements to discourage drinking contaminated waters,⁶⁷ and restrictions on pesticides.⁶⁸

Marketing of Biotech Products, Report of the Panel, Annex, Replies by the Scientific Experts, WT/DS291/R/Add.6 (January 2008) at 170 (providing an example of the technical evidence that the SPS Agreement relies on when defining risk assessment).

62. See Vern R. Walker, *Transforming Science into Law: Transparency and Default Reasoning in International Trade Disputes*, in *RESCUING SCIENCE FROM POLITICS: REGULATION AND THE DISTORTION OF SCIENTIFIC RESEARCH* 165, 167–68 (Wendy Wagner & Rena Steinzor eds., 2006).

63. Agreement on the Application of Sanitary and Phytosanitary Measures, art. 5.1-5.2, Apr. 15, 1994, 1867 U.N.T.S. 493.

64. See IRMGARD MARBOE, *CALCULATION OF COMPENSATION AND DAMAGES IN INTERNATIONAL INVESTMENT LAW* 181 (2009) (“It is increasingly accepted that employing experts usually helps the tribunal better understand the complexities involved in calculating compensation and damages, even though this increases the costs of resolving the dispute and may slow the process.”); José E. Alvarez, *Are International Judges Afraid of Science?: A Comment on Mbengue*, 34 *LOY. L.A. INT’L & COMP. L. REV.* 81, 86 (2011) (“Many, perhaps most, [investor-state cases under ICSID] involving sophisticated investors in complex, ongoing enterprises require heavy expert-laden assessments of fair market or going concern value.”); Joshua B. Simmons, *Valuation in Investor-State Arbitration: Toward a More Exact Science*, 30 *BERKELEY J. INT’L L.* 196, 198 (2012) (“The question of fair market value poses notable challenges for arbitrators because it relates more closely to finance than law. These challenges loom large because arbitrators frequently must determine fair market value in investor-state arbitration.”).

65. See, e.g., *Methanex Corp. v. United States*, Final Award of the Tribunal on Jurisdiction and Merits, ¶ 101 (NAFTA Ch. 11 Arb. Trib. 2005), <http://italaw.com/sites/default/files/case-documents/ita0529.pdf>.

66. See, e.g., *Philip Morris Brands Sàrl v. Oriental Republic of Uruguay*, ICSID Case No. ARB/10/7, Award (2016).

67. See, e.g., *Azurix Corp. v. Argentine Republic*, ICSID Case No. ARB/01/12, Award (2006).

68. See, e.g., *Chemtura Corp. v. Canada*, Award (NAFTA Ch. 11 Arb. Trib. 2010), http://italaw.com/sites/default/files/case-documents/ita0149_0.pdf.

Even when a tribunal or court is reluctant to rely on highly technical evidence, reliance can prove unavoidable. One example is the 2014 ICJ case *Whaling in the Antarctic*, in which Australia argued that Japan had breached an international whaling convention by engaging in a whaling program that permitted lethal sampling of three species of whales.⁶⁹ Japan countered that the lethal sampling fell within an exception in the convention for whaling conducted “for purposes of scientific research.”⁷⁰ The ICJ went to great lengths to avoid expressing an opinion on the scientific merits of the whaling program but nonetheless went on to determine “whether the design and implementation of [the program were] reasonable in relation to achieving [the program’s] stated research objectives.”⁷¹ That led to an analysis of the reasonableness of the program’s lethal sample sizes,⁷² the availability of nonlethal sampling methods,⁷³ and the feasibility of achieving the stated scientific objectives given the sample sizes and methods.⁷⁴ On those questions, the Court could not help but wade into the science and the competing testimonies of the parties’ experts.⁷⁵ In the end, a majority of the Court concluded that the design and implementation of the program was unreasonable in relation to its stated objectives and thus did not fall within the exception for whaling conducted “for purposes of scientific research.”⁷⁶

Partly because of the discomfort with expert evidence on display in *Whaling in the Antarctic*, the ICJ remains a conspicuous outlier in the modern trend toward increased reliance on experts.⁷⁷ Since the 1949

69. *Whaling in the Antarctic* (Austl. v. Japan: N.Z. intervening), Judgment, 2014 I.C.J. Rep. 226, ¶ 48 (Mar. 31).

70. *Id.* ¶ 49.

71. *Id.* ¶¶ 88, 127.

72. *Id.* ¶¶ 181, 193, 195-196.

73. *Id.* ¶ 144.

74. *Id.* ¶ 212.

75. See Daniel Peat, *The Use of Court-Appointed Experts by the International Court of Justice*, 84 BRIT. J. INT’L L. 271, 287-88 (2014) (“[A]lthough the Court [in *Whaling in the Antarctic*] aimed to avoid review of the scientific merits . . . by assessing the reasonableness of sample sizes, sampling method, and other aspects of the programme, it did just that.”); Marco Roscini, *Evidentiary Issues in International Disputes Related to State Responsibility for Cyber Operations*, 50 TEX. INT’L L.J. 233, 263-64 (2015) (“In the *Whaling in the Antarctic* case . . . the experts called by both Australia and Japan gave evidence as expert witnesses and were cross-examined, and the Court relied heavily on their statements to conclude that the special permits granted by Japan for the killing, taking, and treatment of whales had not been granted ‘for purposes of scientific research.’”).

76. *Whaling in the Antarctic* (Austl. v. Japan: N.Z. intervening), Judgment, 2014 I.C.J. Rep. 226, ¶ 227 (Mar. 31).

77. See FOSTER, *supra* note 3, at 121; Joost Pauwelyn, *The Use, Non-Use and Abuse of Economics in WTO and Investment Litigation*, in 43 WTO LITIG., INV. ARB., & COM. ARB., GLOBAL TRADE L. SERIES 172 (Jorge A. Huerta-Goldman, Antoine Romanetti, et al. eds., 2013) (noting the ICJ’s “continued reluctance to engage with expert evidence”); see also *The Gabčíkovo-Nagymaros Project* (Hung. v. Slov.), Merits, 1997 I.C.J. Rep. 7, ¶ 54 (Merits) (Sept. 25) (concluding that the Court need not determine which of the parties’ competing scientific views were more reliable).

Corfu Channel case, the Court has relied on expert evidence in only two cases: the *Gulf of Maine* boundary dispute case in 1984⁷⁸ and the *Whaling in the Antarctic* case in 2014.⁷⁹ Dissenting and separate opinions have long lamented the Court's reluctance to rely on expert evidence,⁸⁰ and the issue of expert evidence—"how to deal with . . . technical and scientific issues"—has been characterized as "one of the biggest problems" facing the Court today.⁸¹

II. PREVAILING PRACTICES FOR EVALUATING EXPERT EVIDENCE

While it is perhaps not possible or worthwhile to evaluate dispute resolution systems on whether they produce outcomes that are "good" or "bad," it is key to the legitimacy of dispute resolution systems that the parties and public have confidence in the design and functioning of the system.⁸² When it comes to expert evidence, international courts and tribunals have done too little to engender that confidence. Rather, they have largely failed to adapt to the reality that experts are often the most critical element of an international dispute, and the increasing reliance on expert evidence in international disputes has continued despite a lack of accepted or consistently applied standards or best practices.⁸³

78. Delimitation of the Maritime Boundary in the Gulf of Maine Area (Can. v. U.S.), Merits, 1984 I.C.J. Rep. 246, 253, 333, 337 (Oct. 12).

79. Whaling in the Antarctic (Austl. v. Japan), 2014 I.C.J. Rep. 148 (Mar. 31), ¶ 20–21. In the *Gulf of Maine* dispute, the Special Agreement between the parties to submit the dispute to the Court specifically required the Court to appoint a technical expert to delineate a precise boundary line, based on the legal principles and rules as set out by the Court. Delimitation of the Maritime Boundary in the Gulf of Maine Area (Can. v. U.S.), Appointment of Expert Order, 1984 I.C.J. Rep. 165, 165–66 (Mar. 30); see *Gulf of Maine*, Merits, 1984 I.C.J. Rep. 246, ¶ 8, 18 (Oct. 12).

80. See Oscar Chinn (U.K. v. Belg.), Merits, 1934 P.C.I.J. (ser. A/B) No. 63, 107–09 (Dec. 12) (individual opinion by Anzilotti, J.); Temple of Preah Vihear (Cambodia v. Thai.), Judgment, 1962 I.C.J. Rep. 6, ¶ 55 (June 15) (dissenting opinion by Koo, J.); Military and Paramilitary Activities in and Against (Nicar. v. U.S.), Judgment, 1986 I.C.J. Rep. 6, ¶¶ 132–134 (June 27) (dissenting opinion by Schwebel, J.); Kasikili/ Sedudu Island (Bots. v. Namib.), Judgment, 1999 I.C.J. Rep. 1045, 1118–19, 1131 (Dec. 13) (separate opinion by Oda, J.); Maritime Delimitation and Territorial Questions Between Qatar and Bahrain (Qatar v. Bahr.), Judgment, 2001 I.C.J. Rep. 40, ¶¶ 39–41 (Mar. 16) (dissenting opinion by Bernárdez, J.); Pulp Mills on the River Uruguay (Arg. v. Uru.), Judgment, 2010 I.C.J. Rep. 14, 108 (Apr. 20) (joint dissenting opinion by Al-Khasawneh, J. and Simma, J.).

81. Riddell, *supra* note 28, at 230. Riddell cites evidence that the Court, without the knowledge of the parties or the public, does "frequently have recourse to expert advice" in the form of temporary staff members of the Court's Registry." *Id.* at 240.

82. See RICHARD A. POSNER, *HOW JUDGES THINK* 3 (2008)

Many of the decisions that constitute the output of a court system cannot be shown to be either 'good' or 'bad,' whether in terms of consequences or of other criteria, so it is natural to ask whether there are grounds for confidence in the design of the institution and in the competence and integrity of the judges who operate it.

83. See, e.g., Singh, *supra* note 42, at 600 ("The reliance upon and widespread usage of expert witnesses [in international criminal cases] has progressed in the absence of clear rules and regulations with regard to the what and how of expert witnesses.").

Existing practices do not provide a sufficient framework to evaluate the reliability of expert evidence. Those practices primarily focus on enhancing the rigor of the adversarial process, appointing “independent” experts, and improving ethical transparency through the disclosure of potential conflicts of interest. Although these practices may sharpen points of disagreement and highlight flagrant biases, and the evaluation of expert evidence has generally improved over time,⁸⁴ none of those practices assist the decision maker in engagement with the *substance* of that evidence. Robust and transparent reliability analysis remains the exception, not the rule.

Although today’s practices in handling expert evidence reflect some minor improvement over nineteenth and early twentieth century practices, they at their core have remained largely unchanged for many years.⁸⁵ It is true, for example, that some early procedures limited parties to written responses to expert testimony.⁸⁶ Yet the ability of a party to question experts at a hearing, which today is enshrined in most procedural rules,⁸⁷ is a longstanding practice, not an innovation prompted by the heightened importance of this sort of evidence. The 1936 statute of the Permanent Court of International Justice, for example, provided for the examination

84. Pauwelyn, *supra* note 77, at 171 (“Ten years later [after the WTO’s report in *EC - Bananas*], the sophistication and quality of analysis in both damage and retaliation calculations have advanced tremendously.”).

85. See FOSTER, *supra* note 3, at 79 (describing “traditional approaches” in international courts and tribunals as including those “that allow the parties to present countervailing scientific evidence, through expert reports, affidavits, advocacy and the appearance and examination of the parties’ own expert witnesses”).

86. See *e.g.* U.S. DEP’T OF STATE, FOREIGN RELATIONS OF THE UNITED STATES 1902: WHALING AND SEALING CLAIMS AGAINST RUSSIA, APPENDIX I, at 428 (1903) (“The Arbitrator observed that the hearing of the experts by the arbitrator could have taken place without the presence of the other party.”); see also *Costa Rica Packet Arbitration* (Gr. Brit. v. Neth.), 184 C.T.S. 240 (1897).

87. See, *e.g.*, IBA RULES ON THE TAKING OF EVIDENCE IN INTERNATIONAL ARBITRATION, art. 6(6) (Int’l Bar Ass’n [IBA], May 29, 2010) (providing that a tribunal-appointed expert “may be questioned . . . on issues raised in his or her Expert Report, the Parties’ submissions or Witness Statement or the Expert Reports made by the Party-Appointed Experts.”); *id.* at art. 8(3)(b) (“With respect to oral testimony at an Evidentiary Hearing . . . following direct testimony, any other Party may question such witness.”); *id.* at art. 8(3)(d) (providing that a tribunal-appointed expert “may be questioned by the Parties or by any Party-Appointed Expert on issues raised in the Tribunal-Appointed Expert Report, in the Parties’ submissions or in the Expert Reports made by the Party-Appointed Experts”); INTERNATIONAL DISPUTE RESOLUTION PROCEDURES, art. 25.4 (Int’l Centre for Dispute Resolution [ICDR], June 1, 2014); INT’L CHAMBER OF COMMERCE RULES OF ARBITRATION, art. 25.4 (Int’l Chamber of Commerce [ICC], Jan. 1, 2012); see also ALAN REDFERN & MARTIN HUNTER, LAW AND PRACTICE OF INTERNATIONAL COMMERCIAL ARBITRATION 312 (4th ed. 2004) (“If each party presents conflicting evidence of technical opinion, the expert witnesses must be prepared to appear in person before the arbitral tribunal for examination. Otherwise, the arbitral tribunal will have no means of evaluating the weight that should be given to the opinions presented by one side or the other.”); ARBITRATION RULES, art. 21.4 (London Court of Int’l Arbitration [LCIA], 2014).

of expert witnesses by the parties and the court⁸⁸ and the same provision remains in force under the 1946 ICJ rules.⁸⁹ The rules governing other twentieth century tribunals and commissions generally permitted parties to examine expert witnesses and comment on their reports.⁹⁰ It has long been true, then, that “the communication of the expert’s report . . . to the parties and the grant of adequate opportunity . . . to comment upon it and to question the expert” are “fundamental procedural safeguards.”⁹¹

As a way to improve upon this longstanding core mechanism for handling expert evidence, commentators have suggested that decision makers should more frequently appoint their own experts, rather than rely on party-appointed experts.⁹² But the authority of a tribunal to appoint its own expert, alone or in addition to any experts that the parties might offer, is nothing new.⁹³ Similarly, experts have long been required to disclose their background, qualifications, and experience, according to a decision maker’s order or other rule.⁹⁴

88. Statute for the Permanent Court of International Justice art. 51, Dec. 16, 1920, 6 L.N.T.S. 407 (1921) [hereinafter P.C.I.J. Statute]; see also WHITE, *supra* note 9, at 39.

89. Statute of the International Court of Justice art. 51, June 26, 1945, 59 Stat. 1062 [hereinafter I.C.J. Statute].

90. See, e.g., Rules of Procedure of the German Arbitral Commission, Rule 45(a), in 1 DECISIONS OF THE ARBITRAL COMMISSION ON PROPERTY, RIGHTS, AND INTEREST IN GERMANY 192, 200 (1958); CHARLES N. BROWER & JASON D. BRUESCHKE, THE IRAN-UNITED STATES CLAIMS TRIBUNAL 201 n.957 (1998) (citing Iran-U.S. Claims Tribunal cases and observing that “[t]he proceedings may involve several phases: the expert’s investigation; submission of a draft report; comments by the parties on the draft report; submission of the expert’s final report; comments by the parties on the final report; and a hearing on the expert’s final report”); Corfu Channel (U.K. v. Alb.), Merits, 1949 I.C.J. Rep. 4, 7 (Apr. 9) (noting that each party cross-examined the expert witnesses of the other party).

91. WHITE, *supra* note 9, at 81 (making this statement in 1965).

92. FOSTER, *supra* note 3, at 79-80

We should endorse a move towards processes for consultation of experts that draw on a blend of investigative and adversarial procedures. . . . A helpful starting point may be to hold an ‘organisational conference’ at the outset of proceedings in order to clarify expectations and requests regarding the procedures to be followed in relation to evidence and proof.

93. See, e.g., Iran-U.S. Claims Tribunal, Final Tribunal Rules of Procedure, art. 27, in 2 IRAN-U.S. Cl. Trib. Rep. 405, 430 (May 3, 1983); G.A. Res. 31/98, Arbitration Rules of the United Nations Commission on International Trade Law [UNCITRAL], art. 27 (Dec. 15, 1976); I.C.J. Statute art. 50, June 26, 1945, 59 Stat. 1062; P.C.I.J. Statute art. 50, Dec. 16, 1920, 6 L.N.T.S. 407 (1921); Convention for the Pacific Settlement of International Disputes art. 90, Oct. 18, 1907, 36 Stat. 2199; see also BROWER & BRUESCHKE, *supra* note 90, at 199-200, 202 (noting that the Iran-U.S. Claims Tribunal, under the UNCITRAL Rules, has the authority to appoint experts, in addition to the experts that the parties may appoint, although the Tribunal “has appointed experts in only eight of the 960 large claims”).

94. See, e.g., Starrett Housing Corp. v. Iran, Final Award No. 314-24-1, 16 IRAN-U.S. Cl. Trib. Rep. 112, 196-97 (1987) (“In determining the weight to be given to the Expert’s Report, the Tribunal must first consider his qualifications. . . . The Tribunal . . . reviewed the Expert’s background and experience before appointing him.”); see also Robert Wisner et al., *Effective Use of Economic Experts in International Arbitration: Counsel’s Role and Perspective*, in EU AND US ANTITRUST ARBITRATION: A HANDBOOK FOR PRACTITIONERS 237, 242-43 (Gordon Blanke & Phillip Landolt eds., 2011).

Other efforts to improve the handling of expert evidence have focused on two procedural aspects: enhancing the disclosure of potential conflicts of interest⁹⁵ and encouraging cooperation and engagement among party-appointed experts with the aim of narrowing the scope of disagreement.⁹⁶ For example, to further witness cooperation the IBA Rules on the Taking of Evidence in International Commercial Arbitration contemplate that tribunals might require the parties' expert witnesses to meet and discuss their reports before the hearing and to develop a joint report summarizing the points of agreement and disagreement.⁹⁷ The IBA Rules also contemplate that decision makers might enhance the adversarial process through "witness conferencing" or "hot-tubbing," in which they may hear more than one expert witness at the same time and encourage the experts to comment on the other experts' views.⁹⁸ Under this approach, party-appointed witnesses present their views and respond to the tribunal's questions.⁹⁹ Another proposed variation of the approach is the "Sachs Protocol," which "envisages at least one 'open' session at which the parties counsel may address and [] question the expert team in the presence of the arbitral tribunal."¹⁰⁰

Existing efforts to improve evidentiary reliability each have their own limitations. Although relying on the adversarial process may assist in exploring and testing the strength of the expert's methods and conclusions,¹⁰¹ it may also involve efforts to muddy rather than clarify the issues

95. See Mark Kantor, *A Code of Conduct for Party-Appointed Experts in International Arbitration – Can One Be Found?*, in 26 *ARB. INT'L* 323, 323 (Park ed., 2010); see also Chartered Inst. of Arb. [CI Arb], *Protocol for the Use of Party Appointed Expert Witnesses in International Arbitration* art. 4(4)(b) (2007) [hereinafter CI Arb Protocol] (requiring experts to disclose "any past or present relationship with any of the Parties, the Arbitral Tribunal, counsel or other representatives of the Parties, other witnesses and any other person or entity involved in the Arbitration").

96. MARBOE, *supra* note 64, at 183 (suggesting that investment tribunals should compare expert reports on damages calculations and "identify those aspects of the valuation where there is agreement and where there is disagreement" or to "request the experts to develop and submit such lists of agreements and disagreements").

97. IBA RULES art. 8(3)(f). The CI Arb Protocol also provides for discussions among party-appointed experts to "identif[y] and list[] the issues upon which they are to provide an opinion," as well as "any tests or analyses which need to be conducted"; "exchange draft outline opinions for the purposes of these meetings"; and prepare a joint statement identifying areas of agreement and areas of disagreement. CI Arb Protocol art. 6(1)(a), (b).

98. IBA RULES art. 8(3)(f); see also Freyer, *supra* note 44, at 439 ("This practice has been used extensively in Asia and Australia and is attracting increasing support among European and American arbitrators.").

99. See *id.*; FOSTER, *supra* note 3, at 123.

100. See Klaus Michael Sachs, *Experts: Neutrals or Advocates*, *INT'L COUNCIL FOR COM. ARB.* (May 25, 2010), http://arbitration-icca.org/conferences-and-congresses/ICCA_RIO_2010/ICCA_RIO_2010_Klaus_Michael_Sachs.html (proposing a variation of expert-witness conferencing); see also Alison Ross, *A Sachs- New Approach to Expert Evidence?*, *GLOBAL ARB. REV.* (May 27, 2010), <http://globalarbitrationreview.com/news/article/28439/a-sachs-new-approach-expert-evidence>.

101. See FOSTER, *supra* note 3, at 79 ("Much can be learnt [through the adversarial process], as avenues of proof are explored and tested by counsel working with their experts.

at hand, or to cast doubt on an expert's testimony through cursory analysis that may have visceral appeal to laymen but lack any technical basis.¹⁰² An expert appointed by a decision maker may do nothing more than act as a de facto member of the court or tribunal whose opinion has decisive effect,¹⁰³ or reflect the strategic demands of the parties on their preferred expert choices.¹⁰⁴ And excessive reliance on credentials can short-circuit a full-fledged examination of expert evidence.¹⁰⁵

More fundamentally, existing efforts may serve to improve the *procedures* used to appoint experts and to evaluate the evidence that they present, but they do not directly address the *substance* of expert evidence. Part III outlines the consequences of failures to engage comprehensively and substantively with expert evidence. In the end, no matter the methods of appointment or the taking of expert evidence, a court or tribunal must still decide for itself whether the testimony is reliable and how much weight to afford it.

III. INADEQUACY OF PREVAILING PRACTICES

Although “[i]t is all too easy to agree that international judges do not handle facts as well as they could,”¹⁰⁶ and many international disputes are resolved in ways that transparently and comprehensively address the reliability of expert evidence, experience continues to demonstrate that, without a comprehensive analytical framework, decision makers often struggle with the complexities of expert evidence. This Part explores the three primary difficulties facing decision makers in international disputes when confronted with expert evidence.

The process of cross-examination is most valuable in testing the strength of propositions relied upon by the parties.”).

102. See FOSTER, *supra* note 3, at 100-01.

103. See MARBOE, *supra* note 64, at 182

[T]he costs of such a tribunal-appointed expert may be considerable. Furthermore, the parties may be concerned that the independent expert will become de facto the fourth, and decisive, arbitrator on the compensation or damages issue. It appears, therefore, that the continental practice of tribunal-appointed expert is only in exceptional cases advisable in international investment disputes.

(internal citations omitted).

104. See Bonneuil & Levidow, *supra* note 49, at 84-86 (observing that, in the WTO *EC — Approval and Marketing of Biotech Products* case, the parties employed different strategies to select experts perceived as favorable to their position, one party by flooding the selection pool with candidates and the other party by proposing only a few candidates and rejecting all of its adversary's candidates, and that the Panel “simply accepted the parties' demands for rejection without evaluating them”).

105. See Wisner, *supra* note 94, at 246.

106. See Alvarez, *supra* note 64, at 96.

A. *Poorly Reasoned Reliance on Expert Evidence*

Experts can play a valuable role in international disputes by gathering and assessing complex technical facts that are outside the general knowledge of the lawyers who generally sit on tribunals and courts.¹⁰⁷ These experts are intended to “assist, educate[,] and advise the arbitral tribunal.”¹⁰⁸ But they may not become the de facto decision makers themselves, because international jurists and arbitrators are generally prohibited from delegating their adjudicatory authority to others.¹⁰⁹ And nothing about specialized testimony makes it inherently reliable or an objective basis for decision making. Instead, the development—and presentation—of expert knowledge is a process susceptible to societal influences, errors, and biases.¹¹⁰ For those reasons, careful reliance on—rather than delegation or outsourcing to—experts is essential to the accuracy and legitimacy of international dispute settlement.

Yet, all too often international courts and tribunals base their decisions on expert evidence without transparently and systematically determining whether that evidence is a reliable basis for decision. This leaves them open to criticisms that they have delegated their decision-making duties to subject matter experts¹¹¹ and elevated expert testimony from the

107. See RIPINSKY & WILLIAMS, *supra* note 43, at 174-75 (“The experts’ function is to assist in ascertaining and assessing facts which involve technical matters outside the general knowledge of judges or lawyers.”); Nicolas Ulmer, *Assessing Damages—Are Arbitrators Good at It? Should They Be Assisted by Experts? Should They Be Entitled to Decide ex aequo et bono? Some War Stories*, 6 J. WORLD INV. & TRADE 11, 11 (2005) [hereinafter *Assessing Damages*] (reasoning that arbitrators “are not generally very good at assessing damages” because they are “mainly lawyers and law professors,” not “accountants,” “investment bankers,” or “quantity surveyors”).

108. See Martin Hunter, *Expert Conferencing and New Methods*, 4 TRANSNAT’L DISP. MGMT. 1, 1 (2007).

109. See WHITE, *supra* note 9, at 13 (“The proper administration of international justice requires that there should be no delegation of the judicial function to anyone outside the tribunal, unless this is clearly the intention of the parties, to cover some special situation.”); MARBOE, *supra* note 64, at 4 (cautioning that, in investment disputes, “the decision about the amount of compensation or damages must not be delegated to the experts”); RIPINSKY & WILLIAMS, *supra* note 43, at 179 (“It is frequently emphasized that tribunals should treat expert opinions, including those produced by the tribunal-appointed experts, as items of evidence and avoid delegating their decision-making powers to experts.”); Caroline E. Foster, *The Consultation of Independent Experts by International Courts and Tribunals in Health and Environment Cases*, 20 FINNISH Y.B. INT’L L. 391, 395 (2009) (recognizing the “need for international courts and tribunals to be alert to the possibility of inadvertent delegation to experts”).

110. See Gruszczynski, *supra* note 32, at 25

The construction of scientific knowledge is a social process that is culturally dependent, and the assessment of risks relies on a number of normative non-scientific approximations and assumptions (*e.g.*[,] owing to gaps in the knowledge, with some of them being irreducible). This means that science is not capable of being a fully objective benchmark in international trade disputes.

111. See WHITE, *supra* note 9, at 13 (noting that the “almost insuperable difficulties of proving that a tribunal had not made an independent evaluation of some, or even all, of the relevant facts, but had followed the evaluation of its expert, can be readily imagined” and

status of evidence to “irrefutable truth.”¹¹² Failure to assess the reliability of expert evidence may result in undue reliance on it, compromising a decision’s legitimacy and inappropriately affecting the regulatory choices of governments, the legitimate commercial interests of companies, or the fate of individuals.¹¹³

The perception of undue reliance on expert evidence has dogged international tribunals and courts. For example, in *Corfu Channel*, the only ICJ case in the pre-modern era to have relied on expert evidence, a dissenting opinion lamented that the Court’s award of damages “makes hardly any reference to the many . . . documents accepted as evidence as damage,” because “something should have been said on their value as evidence,” and the expert’s testimony was accepted without meaningful examination.¹¹⁴ The post-war Franco-Italian Commission, which relied frequently on expert evidence, also seemed to do so at times without any meaningful evaluation of the basis for and quality of expert evidence. For example, a 1963 decision of the Commission stated that, although “[i]t is

citing the 1949 ICJ decision in *Corfu Channel* as one that “rel[ie]d upon the experts’ analysis of the facts to a pronounced degree”); Geoffrey Beresford Hartwell, *Assessing Damages*, 6 J. WORLD INV. & TRADE 7, 8-9 (2005) (“[I]n legal theory—*delegatus non posse delegare*—the tribunal may not delegate its decision-making to the expert. In practice, however, that is difficult to avoid. . . . [the expert] may . . . have decision-making power which is not always apparent. It may be that the tribunal will depend on him very substantially.”); Serge Lazareff, *Assessing Damages*, 6 J. WORLD INV. & TRADE 17, 18 (2005) (“If we turn to the civil-law approach, I must say that I am strongly opposed to the systematic appointment of an expert by the Tribunal because I think it results in delegated justice. It is very difficult for a tribunal not to follow the expert it has itself appointed.”); Ulmer, *supra* note 107, at 12 (noting that arbitrators are very unlikely “to go head-to-head” with certain experts but instead would “very grateful” to leave all the technical issues to them); Alvarez, *supra* note 64, at 97 (“We ought to be equally concerned with whether judges are, in some cases, overly deferential and only too ready to accept some forms of expert-driven ‘fact.’”); Quelques Remarques sur la Preuve devant la Cour Permanente et la Court Internationale de Justice, 7 ANNUAIRE SUISSE DE DROIT INT’L 101 (1950), translated in SANDIFER, EVIDENCE BEFORE INTERNATIONAL TRIBUNALS (1975), at 327 n.144 (“Although the Court, like a national court, is never bound by the conclusions of the experts, these conclusions, as in municipal law, usually exert a relatively large influence upon the decisions taken.”); see also Appellate Body Report, *US — Continued Suspension*, ¶ 436, WTO Doc. WT/DS320/AB/R (adopted Oct. 16, 2008) (“[T]he manner in which [experts’] opinions are solicited and evaluated can have a significant bearing on a panel’s consideration of the evidence and its review of a domestic measure, especially in cases . . . involving highly complex scientific issues.”).

112. Miles, *supra* note 31, at 160.

113. See *id.* (asserting that the complexity of modern international litigation “coupled with a reluctance to engage in in-depth questioning of . . . scientific evidence” may mean that “courts can be overly deferential to scientific expertise or that the weight given to individual written reports could be misplaced”); cf. Alvarez, *supra* note 64, at 97 (criticizing decisions that “rely on a form of ‘scientific’ expertise without exploring whether this is truly warranted . . . silently constrain the regulatory autonomy of governments”).

114. *Corfu Channel* (U.K. v. Alb.), Judgment, 1949 I.C.J. Rep. 252, 253 (Apr. 9) (dissenting opinion by Eèer, J.). The Court compared the damages estimates of tribunal-appointed experts to the claimed damages of the United Kingdom, determining that the latter were reasonable in light of the expert evidence. See *Corfu Channel* (U.K. v. Alb.), Judgment, 1949 I.C.J. Rep. 244, 249 (Dec. 15).

certain that the opinion of the expert does not bind the Commission,” there was no reason not to adopt the expert’s conclusion, “unless his argumentation is in contradiction with the facts of record, with the legal provisions or the rules of logic.”¹¹⁵ The written decision does not indicate what, if any, substantive analysis the Commission undertook in evaluating the expert’s opinion.¹¹⁶

Poorly reasoned evaluations of expert evidence can result not only in the wholesale adoption of an expert’s opinion but also in subtler errors. When decision makers must make numerical estimates, for example, the initial values put forward by the parties and their experts may “anchor” a decision maker’s final calculation—establishing a false equivalence between the competing values of the parties—even if the initial values are extreme and wholly unreasonable.¹¹⁷

When it comes to calculating damages, the anchoring effect has led to a common, although perhaps overly broad, perception that international courts and tribunals do not make fully reasoned judgments when awarding damages but instead “split the baby” and award damages at the midpoint of the two sides’ valuations.¹¹⁸ An oft-cited example is *Santa Elena v. Costa Rica*, where the tribunal expressly chose to value an expropriated asset as exactly half of the sum of the two competing amounts submitted by the parties and their experts, without articulating any analysis of the evidence supporting those amounts.¹¹⁹ The phenomenon may stem from a reluctance to engage with the evidence in a meaningful way, because of

115. Heretier de Sar Mgr. le Duc de Guise, Dec. No. 162 (Nov. 20, 1953), 13 U.N.R.I.A.A. 162, 168 (1963), *translated in* SANDIFER, *EVIDENCE BEFORE INTERNATIONAL TRIBUNALS* (1975), *supra* note 10, at 327.

116. *See id.*

117. *See* Chris Guthrie et al., *Inside the Judicial Mind*, 86 *CORNELL L. REV.* 777, 793-94 (2001) (“[T]he influence on judges of misleading anchors, such as litigants’ requests for damage awards, can produce biased damage awards.”).

118. *See* MARBOE, *supra* note 64, at 182 (“In some cases, the tribunals set the amount of compensation exactly halfway between the claimant’s and the respondent’s valuations, thus appearing to ‘split the baby.’” (internal citations omitted)); RIPINSKY & WILLIAMS, *supra* note 43, at 191 (“[O]n a number of occasions, tribunals appear to have simply ‘split the baby,’ taking the mean between alternative valuations produced by the parties, or were otherwise opaque about how they arrived at the amount of compensation.”); Joost Pauwelyn, *supra* note 77, at 171-72 (“[I]n the not too distant past, WTO arbitrators as well as investor-State tribunals, rather than using sound economic models and data to calculate damages or permitted trade retaliation, were more or less ‘splitting the difference,’ that is adding up what the two opposing parties claimed and then awarding half of that amount.”).

119. *Compañía del Desarrollo de Santa Elena, S.A. v. Republic of Costa Rica*, ICSID Case No. ARB/96/1, Award, ¶¶ 93-95 (Feb. 17, 2000). A number of other decisions, many of them before the last ten years, appear to have made awards essentially at the midpoint of the parties’ estimates, *see, e.g.*, *Middle East Cement Shipping and Handling Co. S.A. v. Arab Republic of Egypt*, ICSID Case No. ARB/99/6, Award, ¶¶ 150-151 (Apr. 12, 2002) (valuing a ship); *Saghi v. Islamic Republic of Iran*, Award, 29 *Iran-U.S. Cl. Trib. Rep.* 20, 55, ¶ 104 (1993) (noting expressly that the award fell between claimants’ and respondent’s estimates).

its complexity and legal decision makers' lack of expertise in financial analytics.¹²⁰

Although most decision makers may not admit to adopting the midpoint between competing valuations,¹²¹ the pull of the less detectable anchoring effect may still influence decisions and lead disputing parties—and their experts—to put forward more extreme valuations as a means to shape a decision maker's damages calculation.¹²² Even decisions that serve as examples of otherwise robust efforts to analyze the reliability of expert evidence have included attempts to reconcile the parties' competing numbers, indicating that the parties may influence the outcome of the dispute simply by raising or lowering their initial values. In *CME v. Czech Republic*, for example, the tribunal stated expressly that it would attempt to “close[]” the “gap between the experts' valuations of roughly USD 210 million” through “a rough assessment.”¹²³

B. *Poorly Reasoned Dismissal of Expert Evidence*

Neglecting to effectively evaluate expert reliability can lead to undue or wholesale adoption of expert opinion, but it can also have quite the opposite effect, including the flippant disregard of evidence that is reliable and that would usefully aid in the decision-making process.¹²⁴ It is too often easier for non-specialists to ignore evidence that they consider too difficult to understand, to cite minor flaws in a technical analysis as a pretext for disregarding it, or to create other doubts as to the reliability of expert evidence, than it is to take the time and effort to understand and appropriately rely on highly technical material.¹²⁵ In other instances,

120. RIPINSKY & WILLIAMS, *supra* note 43, at 122 (noting that reluctance to engage with the details of valuation may lead to “splitting the baby” or some other means of simplifying the valuation process when decision-makers “get lost in the intricacies of valuation techniques”); SIMMONS, *supra* note 64, at 209 (“Because of their backgrounds, arbitrators may be reluctant to immerse themselves in the detailed formulas and spreadsheets submitted by the parties.”).

121. See Christopher R. Drahozal, *Behavioral Analysis of Private Judging*, 67 L. & CONTEMP. PROBS. 105, 118 (2004) (surveying available empirical studies on the anchoring effect—also known as “extremeness aversion”—and concluding that “[a]rbitrators (at least in commercial cases) do not seem to be subject to extremeness aversion”); Richard . Naimark & Stephanie E. Keer, *Arbitrators Do Not “Split the Baby” – Empirical Evidence from International Business Arbitration*, 18 J. INT'L ARB. 573 (2001); see also RIPINSKY & WILLIAMS, *supra* note 43, at 191 (noting that, “in more recent arbitral awards, arbitrators have enhanced the quality of analysis” of expert evidence on valuations).

122. See MARBOE, *supra* note 64, at 182 (reasoning that the “split the baby” problem may “encourage the parties to overestimate or underestimate their claims”); Reed, *supra* note 49, at 5 (citing the “anchoring” effect in cautioning that “[c]ounsel must think carefully the numbers used in estimating damages in initial requests for arbitration”).

123. *CME Czech Republic B.V. v. Czech Republic*, Final Award, ¶ 596 (UNCITRAL Arb. Proc. 2003), <http://italaw.com/sites/default/files/case-documents/ita0180.pdf>.

124. See Miles, *supra* note 31, at 161 (“[P]erhaps[] the risk [in international dispute resolution for a] is that the [expert] evidence is not given sufficient consideration at all.”).

125. See Lawrence, *supra* note 55, at 188 (“Because of their lack of training and expertise in particular disciplines, WTO dispute settlement bodies often fail to adequately assess

“egocentric bias” may mean that decision makers rely on their “own preconceptions and give undue weight to or even seek out evidence to support those preconceptions, rather than fully consider the parties’ arguments.”¹²⁶

Decisions throughout the years have illustrated the difficulties that may result from the poorly reasoned rejection of expert evidence and cursory substitution of a legal decision maker’s lay opinion. Early in its existence, the Iran-U.S. Claims Tribunal, in *Starrett Housing Corp. v. Iran*, appointed a valuation expert, who prepared a detailed report at a cost of over one million dollars, and permitted the parties to comment in writing on the report and examine the witness at a hearing.¹²⁷ Although the Tribunal claimed to rely on the expert’s discounted cash flow (DCF) methodology, it essentially ignored the expert’s application of that methodology to key facts of the dispute.¹²⁸ Instead, the Tribunal made its own qualitative and generic assessment, resulting in significant changes to the amount of damages awarded.¹²⁹ Judge Holtzmann, in a concurring opinion, argued that this represented a departure from an otherwise robust assessment of the expert evidence on valuation.¹³⁰ For one, the Tribunal rejected the expert’s use of compound interest,¹³¹ despite the “modern economic reality” confirmed by the expert, that investors were required to pay on the loans that had financed the expropriated project.¹³²

The 1989 decision in the *Salmon and Herring v. Canada* case, which arose under the United States-Canada Free Trade Agreement, has been criticized for the tribunal’s apparent “willingness to substitute [its] own judgment for the numerical determinations of governmental experts based on [its] own reading of scientific texts.”¹³³ The United States had chal-

and incorporate scientific contributions.”); Ulmer, *supra* note 107, at 14 (recounting from an experience as counsel in a commercial arbitration that “it is, in general, much easier to pick holes, to create doubts, in the damages presentation to actually put forward an iron-clad one”).

126. See Reed, *supra* note 49, at 7.

127. See *Starrett Housing Corp. v. Iran*, Final Award, 16 Iran-U.S. Cl. Trib. Rep. 112, 120, ¶ 11 (1987).

128. *Id.* at 117, ¶¶ 337-343 (1987); see generally RIPINSKY & WILLIAMS, *supra* note 43, at 176 n.270 (observing that the extensive expert evidence in *Starrett Housing* went “largely ignored” by the tribunal).

129. *Starrett Housing*, 16 Iran-U.S. Cl. Trib. Rep. at 117, ¶¶ 337-343.

130. *Starrett Housing*, 16 Iran-U.S. Cl. Trib. Rep. at 241-54 (1987) (Holtzmann, J., concurring opinion).

131. See *Starrett Housing*, 16 Iran-U.S. Cl. Trib. Rep. ¶ 370.

132. *Starrett Housing*, 16 Iran-U.S. Cl. Trib. Rep. at 249-54 (Holtzmann, J., concurring opinion).

133. See, e.g., David A. Wirth, *The Role of Science in the Uruguay Round and NAFTA Trade Disciplines*, 27 CORNELL INT’L L.J. 817, 845 (1994). Although the “experts” in this case were not testifying witnesses—they were specialists within the Canadian regulatory agencies responsible for the challenged regulation—the panel’s consideration of their purported expertise is instructive of the potential for decision makers to dismiss expert evidence without adequately assessing the reliability of that evidence.

lenged Canadian regulations that required commercial fishing companies to off-load on Canadian territory one hundred percent of all caught herring and salmon, including fish immediately destined for export.¹³⁴ The panel did not express a conclusion that the Canadian government specialists were not experts in their field, nor did the panel rely on the countervailing testimony of other experts. Yet the panel reviewed statistics and fisheries texts on its own and “on the basis of logical analysis” concluded that Canada’s regulation did not fall within the exception for measures “relating to the conservation of exhaustible natural resources” because “reliable sampling data can be obtained without requiring access to 100% of the catch.”¹³⁵

Tribunals and courts have been criticized consistently over the years for dismissing expert evidence without adequate reasoning when tasked with calculating damages. Many legal decision makers have been reticent to use “income-based” valuation methods, such as the DCF method, that are the most widely used in actual business practice,¹³⁶ although there is some indication that the method is increasingly accepted in international disputes, at least in concept.¹³⁷

Even when a tribunal accepts the methodology as a general matter, difficulty remains in applying the methodology to the specifics of individual disputes. In the 2014 decision in *Venezuela Holdings v. Venezuela*, for example, the tribunal’s valuation analysis focused on whether the “discount rate,” which translates future cash flows into a present value, should account for the risk that the host country will confiscate the investment.¹³⁸

134. *In the Matter of Canada’s Landing Requirement for Pacific Coast Salmon and Herring*, U.S.-Canada Free Trade Agreement Binational Panel Review, Panel No. CDA-89-1807-01 (Oct. 16, 1989).

135. *Id.* ¶¶ 7.08, 7.21, 7.29.

136. *See* MARBOE, *supra* note 64, at 206–07.

[I]n the practice of international courts and tribunals, the income approach has long been regarded with skepticism. The necessary forecasts as well as the discounting process appeared to be hardly understandable for lawyers who are not specially trained in economic disciplines. . . . This reluctance to apply an income-oriented valuation approach in international proceedings has been predominant for decades despite the fact that the income approach is the most widely used approach in economic reality and, thus, routinely applied by valuation experts in many different contexts.

MARK KANTOR, VALUATION FOR ARBITRATION 53 (2008) (“The contrast between the preference among international investment tribunals for Asset-Based methods and the preference among valuation organizations and professions for Income-Based methods is striking.”).

137. *See* MARBOE, *supra* note 64, at 215.

138. *Venezuela Holdings B.V. v. Bolivarian Republic of Venezuela*, ICSID Case No. ARB/07/27, Award, ¶¶ 360–368 (2015). *See generally* RICHARD A. BREALEY ET AL., PRINCIPLES OF CORPORATE FINANCE 16 (8th ed. 2006) (“To calculate present value, we discount expected payoffs by the rate of return offered by equivalent investment alternatives in the capital market. This rate of return is the discount rate, hurdle rate, or opportunity cost of capital.”). Future cash flows are worth less than cash flows today. *See id.* (“The first basic principle of finance is that a dollar today is worth more than a dollar tomorrow, because the

After determining that any valuation must account for the risk that Venezuela would confiscate the investment, the tribunal concluded that it was “unable to adopt the approach used by the Claimants’ expert, which does not take this risk into account.”¹³⁹ Venezuela’s experts had taken the confiscation risk into account, calculating a range of discount rates more than twice the rate offered by the claimants’ expert and greatly reducing the overall valuation.¹⁴⁰ Adopting a discount rate within that range, the tribunal noted only that other tribunals “in circumstances comparable to the present case” had adopted similar discount rates, including a commercial arbitration tribunal evaluating contract claims related to the same project.¹⁴¹ The tribunal did not evaluate the underpinnings of Venezuela’s expert evidence whatsoever, nor did it explain why it considered the decisions of other tribunals—all but one considering claims against countries other than Venezuela—were “in circumstances comparable” to this particular time period, industry, and economy.¹⁴²

A larger debate—about whether to ever rely on expert evidence—has been playing out prominently in the ICJ in recent years. Those opposed to reliance on expert evidence have been in the majority in most recent cases but dissenting voices have grown louder and more numerous over the years.¹⁴³ A trio of cases illustrates the criticisms the Court has received

dollar today can be invested to start earning interest immediately. Financial managers refer to this as the time value of money.”).

139. Venezuela Holdings B.V., ICSID Case No. ARB/07/27, Award, ¶ 365.

140. See *id.* ¶ 366.

141. See *id.* ¶¶ 367–368.

142. See *id.*

143. See, e.g., Oscar Chinn (U.K. v. Belgium), Judgment, PCIJ Rep. Series A/B No. 63, 107-09 (individual opinion of Judge Anzilotti); Temple of Preah Vihear (Cambodia v. Thailand) (Merits), 1962 I.C.J. Rep. 6, ¶ 55 (dissenting opinion of Judge Wellington Koo); Military and Paramilitary Activities in and Against (Nicaragua v. U.S.) (Merits), 1986 I.C.J. Rep. 6, ¶¶ 132–134 (dissenting opinion of Judge Schwebel); Kasikili/Seduku Island (Botswana v. Namibia) (Merits), 1999 I.C.J. Rep. 1045, 1118–19 (separate opinion of Judge Oda); Maritime Delimitation and Territorial Questions (Qatar v. Bahrain) (Merits), 2001 I.C.J. Rep. 40, 275 (dissenting opinion of Judge Torres Bernárdez); Pulp Mills on the River Uruguay (Argentina v. Uruguay) (Merits), 2010 I.C.J. Rep. 14, 108. In the first of these cases, the P.C.I.J. in the *Oscar Chinn* case declined—against significant dissent—to appoint an expert to assist in determining whether Belgium had established a de facto monopoly in the Belgian Congo, in violation of its international obligations to the United Kingdom. See Oscar Chinn (U.K. v. Belgium), Judgment, P.C.I.J., ser. A/B, No. 63, at 88 (1934) (reasoning that the details of the economic conditions related to the claims could not change the legal conclusion that Belgium had not violated its international obligations); *id.* at 108 (individual opinion of Judge Anzilotti) (“[T]his necessitates the appraisal of a number of technical considerations which only experts in fluvial navigation and in the economic conditions of the Congo could adequately present to the Court.”); see also SANDIFER, EVIDENCE BEFORE INTERNATIONAL TRIBUNALS, *supra* note 10, at 334–35 (1975) (criticizing the *Oscar Chinn* majority’s refusal to appoint an expert); Thomas M. Franck, *Fact-Finding in the I.C.J.*, in FACT-FINDING BEFORE INTERNATIONAL TRIBUNALS 21, 21 (1992) (faulting the Court for resolving “key factual issues” in *Temple of Preah Vihear*, the Nicaragua Case, and the Advisory Opinion in *Western Sahara* “without leaving The Hague to take testimony, appointing masters, or otherwise familiarizing itself with the sights, sounds and smells of the place in which the facts were embedded.”).

for refusing to consider expert evidence. The first case, *Maritime Delimitation and Territorial Questions Between Qatar and Bahrain*, centered on whether the landmass Fasht Al Azm and Sitrah Island were joined, an issue that affected the extent of each country's maritime areas.¹⁴⁴ Qatar argued that a navigable channel had always separated the two until Bahrain filled it in 1982, and the parties submitted competing expert reports on the existence of a permanently navigable channel before that time.¹⁴⁵ The Court viewed the expert evidence as too difficult to evaluate and based its decision on other issues without reaching a conclusion on the pre-1982 existence of a channel.¹⁴⁶ Others have noted that the Court-determined "maritime" boundary inadvertently crosses over dry land, a situation that would have likely not occurred had the Court relied on expert assistance.¹⁴⁷

In its 1997 *Gabèikovo-Nagymaros Project* decision, the ICJ essentially refused to consider the parties' expert evidence¹⁴⁸ and this decision has been cited as an example of the ICJ's inability or unwillingness to deal with technical evidence.¹⁴⁹ At issue in *Gabèikovo-Nagymaros* was the agreement by Hungary and Slovakia in a 1977 treaty to jointly build a system of hydroelectric-generating locks on the Danube River.¹⁵⁰ The dispute arose after Hungary, facing public criticism of the financial benefits and environmental impact of the project, halted construction of the project, citing a "state of ecological necessity."¹⁵¹

At the ICJ, both parties submitted expert reports on the ecological effects of the locks, with Hungary arguing that they would cause eutrophication in the reservoir, a build-up of silt within the dredged channel, and negative effects on the plants and animals that would be separated from the main river channel.¹⁵² In the Court's view, however, it was "not necessary . . . to determine which of those points of view is scientifically better founded" because the environmental risks were at most long-term and uncertain, abandonment of the project was not the only mitigating measure Hungary could have taken, and Hungary had helped create the situation when it concluded the 1977 treaty.¹⁵³

A similar pattern was repeated in 2010, in the *Pulp Mills on the River Uruguay* case. The case readily lent itself to technical evidence, as it cen-

144. *Maritime Delimitation and Territorial Questions Between Qatar and Bahrain* (Qatar v. Bahrain), Judgment, 2001 I.C.J. Rep. 40, ¶¶ 217–218 (2001).

145. *See id.* ¶ 189.

146. *See id.* ¶¶ 190, 218.

147. *See, e.g.,* Riddell, *supra* note 28, at 243.

148. *See Gabèikovo-Nagymaros Project* (Hungary v. Slovakia), Judgment, 1997 I.C.J. Rep. 7, 42 (1997).

149. ANNA RIDDELL & BRENDAN PLANT, EVIDENCE BEFORE THE INTERNATIONAL COURT OF JUSTICE 348 (2009).

150. *Gabèikovo-Nagymaros*, Judgment, 1997 I.C.J. Rep. at 23.

151. *Id.* at 35.

152. *Id.*

153. *Id.* at 42, 55, 57.

tered on Argentina's claim that a Uruguayan plan to build pulp mills on the River Uruguay, the boundary between Uruguay and Argentina, would significantly pollute the river, in breach of a 1975 treaty between the two countries.¹⁵⁴ Even though the parties produced "a vast amount of factual and scientific material,"¹⁵⁵ the Court did not discuss the merits of the extensive expert evidence and expressly stated that it would not address the reliability of that evidence or the weight accorded to it.¹⁵⁶ In dissent, Judges Al-Khasawneh and Simma lamented the Court's approach, reasoning that "[t]he Court on its own is not in a position adequately to assess and weigh complex scientific evidence of the type presented by the Parties."¹⁵⁷ Instead, "[t]he adjudication of disputes in which the assessment of scientific questions by experts is indispensable, as is the case here, requires an interweaving of legal process with knowledge and expertise that can only be drawn from experts properly trained to evaluate the increasingly complex nature of the facts put before the Court."¹⁵⁸

C. *Use of Expert Evidence as a Veneer to Enhance Legitimacy*

Given the amount of time and paper devoted to expert evidence and the potential influence that it may have on the outcome of international disputes, we would expect that the written decisions resolving those disputes would at the very least explain whether and how the decision-makers relied on expert evidence.¹⁵⁹ But too many decisions do not do that. Failure to record in written decisions the extent of reliance on expert evidence compounds the potential for undue reliance on or unreasoned rejection of expert evidence. As a dissent in the 1949 ICJ *Corfu Channel Case*

154. Pulp Mills on the River Uruguay (Arg. V. Uru.), Judgment, 2010 I.C.J. Rep. 14, ¶¶ 1, 22 (Apr. 20).

155. *Id.* ¶ 165; see also Bruno Simma, *The International Court of Justice and Scientific Expertise*, 106 AM. SOC'Y INT'L L. 230, 230 (2012) (noting that both parties in the *Pulp Mills* case produced "large amounts of scientific expert evidence, both in their written pleadings and in the oral hearings").

156. Pulp Mills on the River Uruguay, 2010 I.C.J. Rep. 14, ¶¶ 166, 168 ("[T]he Court does not find it necessary in order to adjudicate the present case to enter into a general discussion on the relative merits, reliability and authority of the documents and studies prepared by the experts and consultants of the Parties."); see also *id.* ¶ 236 ("[I]n assessing the probative value of the evidence placed before it, the Court will principally weigh and evaluate the data, rather than the conflicting interpretations given to it by the Parties or their experts and consultants."); FOSTER, *supra* note 3, at 93 (noting that the ICJ, in *Pulp Mills*, "did not find it necessary to enter into a general discussion on their merits, reliability[,] and authority.").

157. Pulp Mills on the River Uruguay (Arg. v. Uru.), 2010 I.C.J. Rep. 108, ¶ 4 (Apr. 20) (Al-Khasawneh, J. and Simma, J., joint dissenting opinion).

158. *Id.* ¶ 3.

159. Foster, *supra* note 45, at 161 ("We can expect judgments to be written . . . in terms reflecting how the [ICJ] has made use of expert opinion. Transparency in the reasoning of judgments will help them to stand up to scrutiny, including where expert opinion has fed into the more legal and interpretive tasks required of the Court."); Pauwelyn, *supra* note 77, at 189 ("[I]t is equally important that the judge when weighing the evidence sufficiently explains his or her thought process and engages with the details of each study.").

reasoned, “[a]ccording to a quite general rule of procedure, the Court is not bound by the opinion of experts [and] [t]he Court may reject or accept it[,] but it must always give sufficient reasons.”¹⁶⁰

Others have argued that decision makers often marshal enormous amounts of expert evidence as part of the decision-making process, or in their written decisions, as a way to enhance their perceived legitimacy,¹⁶¹ even in venues, such as the WTO, that routinely seek input from experts.¹⁶² From this perspective, a decision maker may cite to expert evidence to demonstrate that it has left no stone unturned, without engaging with the evidence to better understand the substance of the issues in dispute or to assess its quality.

Expert evidence that complicates rather than quickly confirms other aspects of a decision maker’s analysis may create incentives to ignore the evidence in the written decision, but the reticence to make meaningful use

160. Corfu Channel (U.K. v. Alb.), 1949 I.C.J. Rep. 4, 253 (Dec. 15) (Eèer, J., dissenting opinion).

161. See Bonneuil & Levidow, *supra* note 49, at 97 (criticizing WTO decisions under the SPS Agreement for failing to engage with scientific evidence and framing that evidence” in a way that allows WTO judges to avoid any explicit engagement with scientific knowledge,” so that they can use that evidence as an legitimizing “imprimatur” for “its own judgements on substantive scientific issues”); see also Lazareff, *supra* note 111, at 17 (“Assessing damages is the *parent pauvre* of arbitration, the neglected aspect. It is almost in the context of arbitration, the midnight clause of a contract, and it is very distressing to read in so many awards that ‘the Tribunal, having at its disposal all the elements of the case, orders A to pay B US \$140 million.’”); Simma, *supra* note 155, at 230 (noting that the parties in the *Pulp Mills* case, see *supra* note 156, “produced large amounts of scientific expert evidence” and explaining that as an ICJ judge in the case the author, when “confronted with the mass of such highly complex technical material” did “not feel[] capable of drawing the necessary legal conclusions”).

162. See FOSTER, *supra* note 3, at 133 (“[E]ven in the WTO, the full impact of expert evidence is not always apparent on the face of the report. For example in *United States – Import Prohibition of Certain Shrimp and Shrimp Products*, the Appellate Body’s final disposal of the case was undoubtedly informed by the experts’ advice that sea turtles faced different problems in different locations, but in reaching this decision the Appellate Body did not directly rely on the panel’s consultation of scientific experts.”); cf. Walker, *supra* note 62, at 300–02 (“The [EC —]Hormones Panel took the worst approach [in evaluating whether a measure was “based on” a risk assessment in accordance with the SPS Agreement]: (1) pretend not to pass judgment on the merits of past scientific reports while implicitly agreeing with their conclusions, and (2) formally preclude rebuttal explanation in the WTO proceeding by members who disagree with those scientific reports, saying that WTO panels do not themselves conduct risk assessments.” (citation omitted)); *id.* at 310

[T]he [EC — Hormones] Panel was attempting an impossible task—making determinations about risk assessment documents without actually understanding or taking a position on risk assessment issues. . . . The Hormones Panel should have used all of the relevant scientific information made available to it by the parties and by its own appointed experts, but should have used it for the limited purposes of finding whether there was any reasonable scientific basis for the risk determinations of the European Communities.

(footnote omitted).

of expert evidence can be so strong that decision-makers will not rely on it even when it appears to support their conclusions.¹⁶³

IV. MORE ROBUST ASSESSMENTS OF EXPERT RELIABILITY

How can legal decision makers address these challenges and enhance their handling of expert evidence? Part II of this Article outlined recent proposals to improve the use of expert evidence by international courts and tribunals through procedural and transparency mechanisms, including enhanced conflict-of-interest disclosure requirements and hearings that emphasize joint evaluations of competing experts. While these mechanisms may prove valuable, they are not enough. As others have recognized, decision makers cannot engage passively with complexity, particularly when it comes to expert evidence.¹⁶⁴ From the perspective of ICJ Judges Al-Khasawneh and Simma, “the traditional methods of evaluating evidence are deficient in assessing the relevance of such complex, technical and scientific facts.”¹⁶⁵ Decision makers must recognize the “strengths and weaknesses of expert opinions,” meaningfully distinguish among competing opinions and methodologies, and understand whether an expert’s methodology can be reliably applied to the specific circumstances of the case at hand.¹⁶⁶ Procedures and transparency alone cannot

163. See Jürgen Kurtz, *Science as a Common Proxy for Rational Regulation Across International Trade and Investment Law*, in *SCIENCE AND TECHNOLOGY IN INTERNATIONAL ECONOMIC LAW: BALANCING COMPETING INTERESTS* 134, 146 (Bryan Mercurio & Kuei-Jung Ni eds., 2014) (noting that the *Methanex* tribunal’s “careful and extensive assessment of the environmental justifications for the MTBE ban” based on “significant attention to extensive witness testimony” “receives no attention whatsoever in the Tribunal’s legal analysis of fair and equitable treatment” but “is instead quarantined in the factual findings of the award,” even though “this review . . . provides clear and compelling evidence, as the Tribunal itself determines, of a clearly rational and science-based approach to risk regulation”).

164. See *MARBOE*, *supra* note 64, at 4 (reasoning, in the context of damages calculations in investment disputes, that “[i]n order to assess the figures submitted by experts, a basic understanding of the methods and a willingness to consider them is of vital importance” to avoid delegating “the decision about the amount of compensation or damages” to the experts); *RIPINSKY & WILLIAMS*, *supra* note 43, at 180 (noting in the valuation context that “arbitrators may need to make diligent efforts to understand economic and valuation issues on their own” to be able “to make use of experts’ reports and at the same time exercise their decision-making power”); *Foster*, *supra* note 45, at 144 (“An international court must be ready to determine objective compliance with legal tests in order to enforce the balance of interests they represent. In order to achieve this, the Court [ICJ] will need to engage in the science, rather than to be viewed as shying away from doing so.”); *id.* (“The Court [ICJ] will need to ensure it has the capacity to achieve the sufficiently reliable insights into the science necessary for a sound resolution of both disputes, taking into account also the interests of the wide international community.”).

165. *Pulp Mills on the River Uruguay (Arg. v. Uru.)*, 2010 I.C.J. Rep. 108, ¶ 3 (Apr. 20) (Al-Khasawneh, J. and Simma, J., joint dissenting opinion).

166. See *RIPINSKY & WILLIAMS*, *supra* note 43, at 190–91

[L]awyers need to have an idea about the basic workings of valuation techniques in order to recognize the strengths and weaknesses of expert opinions, be able to discern the ‘drivers’ of valuation, understand why different valuation techniques may yield widely divergent results and, indeed, why even the same valuation

ensure that expert evidence is reliable and that decision makers reliably understand and apply that evidence in crafting legal decisions. As Lucy Reed has reasoned in the context of investor-state arbitration, decision-makers “can do a better job” in fact-finding “by following basic practices to control and mitigate complexities.”¹⁶⁷

When it comes to mitigating the complexities of expert evidence, the focus should be on the reliability of that evidence: the reliability of the methodology and the reliability of that methodology’s application to the facts of the case. The core of this Article’s proposal, then, is that international adjudicators should adopt an analytical framework for determining whether an expert’s testimony is reliable, apply that framework to each expert, clearly set out in the written decision the framework and how it was applied, and weigh the evidence in accordance with its assessed reliability.

A. Applying a Reliability Framework

The core of this Article’s proposal is an illustrative and flexible checklist for evaluating the reliability of expert evidence. In other fields, as “[k]now-how and sophistication have increased remarkably,”¹⁶⁸ decision-makers have turned to checklists to “correctly” and “reliably” deliver the benefits of that know-how and sophistication.¹⁶⁹ Engineers, doctors, and pilots—to name just a few professions—use them to remind themselves of the “minimum necessary steps” and to “establish[] a higher standard of baseline performance.”¹⁷⁰ Checklists, then, can “help even experienced professionals” to “slow down and take account of all relevant factors”¹⁷¹ and to ensure that decision-makers have received all information critical to their decision.¹⁷² And because checklists also set out the analytical steps for all to see, they can lead to more transparent decision-making.¹⁷³

method applied by different experts can lead to very different results. This knowledge is necessary for an informed and considered decision.

See also Walker, *supra* note 62, at 169 (“[T]he SPS Agreement’s approach of using scientific research as a neutral arbiter over the legitimacy of sanitary and phytosanitary measures depends upon the WTO’s being able to distinguish legitimate or sound science from an inadequate risk assessment.”).

167. Lucy Reed, *Confronting Complexities in Fact-Finding and the Nature of Investor-State Arbitration*, in AM. SOC’Y INT’L L., PROCEEDINGS OF THE 106TH ANNUAL MEETING 233, 233 (2013).

168. ATUL GAWANDE, THE CHECKLIST MANIFESTO 11 (2009).

169. *Id.* at 13.

170. *Id.* at 39.

171. Reed, *supra* note 49, at 14 (“One thing we see in heuristics studies is the value of checklists. Checklists help even experienced professionals—for example, pilots and surgeons—to slow down and take account of all relevant factors.”).

172. GAWANDE, *supra* note 168, at 79, 99, 101.

173. *Id.* at 79, 99, 101.

When considering the “evidential value of any evidence produced . . . under all the circumstances” of a highly complex international dispute,¹⁷⁴ an expert reliability checklist could perform similar functions. The result of the checklist would be a clearer analytical roadmap—clearer to the disputing parties, the public, and future decision makers—that a tribunal can use to systematically weigh expert testimony and ultimately the merits of the parties’ arguments.¹⁷⁵ More reliable and relevant testimony would be entitled, as it is now, to greater weight than less reliable and relevant testimony.¹⁷⁶

At the same time, any proposal for evidence-taking in international law should rely on insights developed within international dispute resolution systems and avoid disrupting the essential elements of those systems. Because international law recognizes no hard-and-fast rules for assessing the reliability of expert evidence, leaving adjudicators free to assess and weigh evidence as they see fit,¹⁷⁷ this proposal does not seek to limit the flexibility or discretion inherent in international adjudication. A checklist

174. *Shufeldt (U.S. v. Guat.)*, 2 R.I.A.A. 1079, 1083 (1930) (“[I]t is clear that international courts are by no means as strict as municipal courts and can not [sic] be bound by municipal rules in the receipt and admission of evidence. The evidential value of any evidence produced is for the international tribunal to decide under all the circumstances of the case.”).

175. *See Pauwelyn, supra* note 77, at 185 (suggesting that adjudicators adhere to, “openly state,” and “enforce” a set of “best practices” based on European competition law in evaluating quantitative economic data).

176. *See, e.g., Corfu Channel (U.K. v. Alb.)*, Judgment, 1949 I.C.J. Rep. 4, 21 (Dec. 15) (stating that the Court “cannot fail to give great weight to the opinion of the [e]xperts who examined [the evidence] in a manner giving every guarantee of correct and impartial information.”); U.S. DEP’T OF STATE, *supra* note 86, at 428–29 (noting that while “[i]n an arbitration between States it is of far greater interest than in purely juridical proceedings to draw forth all evidence, whether direct or indirect, which may serve to give full light,” the arbitrator would take into account the circumstances of the expert testimony when weighing the evidence); *see also* Michael Straus, *The Practice of the Iran-U.S. Claims Tribunal in Receiving Evidence from Parties and from Experts*, 3 J. INT’L ARB. 57, 58 (1986) (“[A]ll testimony is ultimately subject to the arbitrators’ judgment as to its probative value based on such factors as credibility and relevance.”); JEFF WAICYMER, *PROCEDURE AND EVIDENCE IN INTERNATIONAL ARBITRATION* 943 (2012) (“In most cases, the [arbitral] tribunal will simply allow the parties to present the experts of their choice and allow submissions as to relative expertise to go to questions of weight.”).

177. *See* GARY BORN, *INTERNATIONAL ARBITRATION: CASES AND MATERIALS* 767 (2015) (“[A]rbitrators in interstate arbitrations enjoy broad discretion to make evidentiary decisions on subjects such as admissibility, weight, and credibility of evidence, usually without reference to municipal evidentiary rules.”); FOSTER, *supra* note 3, at 178 (noting that international law recognizes no “rules requiring that weight will only be placed on scientific evidence that has been subject to peer review and publication”); DURWARD V. SANDIFER, *EVIDENCE BEFORE INTERNATIONAL TRIBUNALS* 2 (1939) (“[I]nternational tribunals are . . . intolérant of any restrictive rules of evidence which might tend to confine the scope of a search after those facts.”); J. L. SIMPSON & HAZEL FOX, *INTERNATIONAL ARBITRATION: LAW AND PRACTICE* 192 (1959) (“In international law there are no general rules requiring the exclusion of categories of evidence.”); *see also* WHITE, *supra* note 9, at 7 (quoting Sandifer’s assertion with approval).

is not a formula,¹⁷⁸ and decision makers would of course remain free to tailor the checklist—add, subtract, or modify elements—to fit particular circumstances of a dispute. Procedural and evidentiary procedures would thus remain flexible to adapt to the particular circumstances of each case.¹⁷⁹

]B. *A Proposed Checklist to Evaluate Expert Reliability*

This Article sets out six suggested framing questions—drawn from international law decisions, practice, and commentary—that might usefully serve as a starting point for evaluating the reliability of expert evidence. These questions emphasize the reliability of the *process* by which an expert has developed and presented evidence, not on the validity of the *outcome* of that process per se. The proposed checklist seeks to focus the inquiry on proxies that are most accessible to lawyers and judges—that is, most like the ways in which they examine and weigh other types of evidence—and does not demand that a decision maker become the arbiter of competing claims of truth in technical fields where she lacks the necessary competence and training.¹⁸⁰

Importantly, when presented with expert evidence, a decision maker ought to record in the written decision whether, how, and why it has relied on that evidence. A record of the question presented—including applicable legal and evidentiary rules—and of the extent to which a decision maker relied on expert evidence in answering that question is essential to avoiding any doubt about the relative roles of expert and legal decision maker.¹⁸¹ Without expressly connecting the question presented, the ex-

178. GAWANDE, *supra* note 168, at 167.

179. Because “[e]ach arbitration presents its own peculiar problems,” “[r]ules that may function well enough in one arbitration may lead to a complete breakdown when introduced in another arbitration possessing totally different characteristics.” KENNETH S. CARLSTON, *THE PROCESS OF INTERNATIONAL ARBITRATION* 30–31 (1946). “[P]rocedural rules should” thus “be carefully adapted to the requirements of each arbitration as it arises so that it may be consummated speedily, economically and justly.” *Id.* at 4. “[T]he type of the procedure can always be adjusted to the complexity and volume of the litigation to be submitted for decision.” *Id.* at 260. Domestic evidentiary rules also do not prescribe strict rules that courts must follow when examining the reliability of experts. The U.S. Supreme Court, for example, has recognized that “the factors identified in *Daubert [v. Merrell Dow Pharmaceuticals, Inc.]* may or may not be pertinent in assessing reliability, depending on the nature of the issue, the expert’s particular expertise, and the subject of his testimony.” See *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 150 (1999).

180. See Marcos A. Orellana, *The Role of Science in Investment Arbitrations Concerning Public Health and the Environment*, 17 Y.B. INT’L ENVTL. L. 48, 49 (2006) (“[E]mphasis on the process by which the scientific enterprise has been conducted, rather than on absolute scientific validities, may prevent arbitrators from becoming ensnared in deciding over the truth of scientific claims, especially in situations of scientific uncertainty.”).

181. See Francesca Romanin Jacur, *Remarks on the Role of Ex Curia Scientific Experts in International Environmental Disputes*, in *INTERNATIONAL COURTS AND THE DEVELOPMENT OF INTERNATIONAL LAW* 441, 452 (Nerina Boschiero et al. eds., 2013) (“Transparency should also be reflected in the final judgment or award. The reasoning of the judiciary should be spelled out clearly, showing how certain conclusions were reached and the contribution of the experts.”).

pert evidence, and the decision, a court or tribunal may leave itself open to criticisms, including those outlined in Part III.

1. Will expert evidence usefully aid in answering the question presented?

The threshold question in determining the weight to accord expert testimony should be whether the testimony is relevant to the question presented and will aid in the decision-making process. If the issue can be decided on the basis of other evidence—for example, documentary evidence and fact witness testimony—or if the question requires only a legal determination, a decision maker may have no need to rely on an expert’s testimony.¹⁸² Asking this threshold question is important because, although expert testimony can assist in clarifying an adjudicator’s decision, it can also confuse a question that would have been otherwise easily and transparently decided without the expert’s input. Answering the question—in writing—also avoids any doubt as to whether, how, and why a decision-maker has relied on expert evidence in reaching its decision.

Today, arbitrators and courts are broadly empowered to consider and rely on expert testimony as they see fit.¹⁸³ But of course just because a tribunal *can* rely on expert testimony does not mean that it *should*. The power to consider expert testimony must include the power to afford only the weight that the testimony deserves. This is a natural inquiry in civil law jurisdictions and in international adjudicatory systems where the tribunal—not the parties—decides whether to appoint experts.

What is the technical question presented?—As simple as it may seem, it is essential to begin with a clear sense of the question that is presented, based on the applicable legal and evidentiary rules, and the factual circumstances. What was the value of an investment at the time of its expropria-

182. See *Pulp Mills on the River Uruguay (Arg. v. Uru.)*, 2010 I.C.J. Rep. 121, ¶ 11 (Apr. 20) (Keith, J., concurring opinion) (“[R]esponsibility of making decisions on the matters of scientific dispute arises only if the matters require decision in the course of the Court determining whether or not Argentina had made out its claim.”); *Factory at Chorzów (Ger. v. Pol.)*, Judgment, 1928 P.C.I.J. (ser. A) No. 17, 49, 51 (Sept. 13) (noting that the Court was not “satisfied with the data for assessment supplied by the Parties” with respect to compensation owed to the German government and providing for the appointment of experts “to obtain further enlightenment in the matter”); *Starrett Housing Corp. v. Iran*, Final Award, 16 Iran-U.S. Cl. Trib. Rep. 112, ¶ 264 (1987) (noting that, although it is not always necessary to appoint valuation experts where “there is sufficiently clear evidence on which to base a decision on compensation,” tribunals may look to expert evidence “to obtain any technical information that might guide it in the search for the truth” (partially quoting *Corfu Channel (U.K. v. Alb.)*, Merits, 1949 I.C.J. Rep. 4, 20 (Dec. 15))); see also Dave, *supra* note 28, at 814 (“It is important to bear in mind that if direct evidence is available and is acceptable, it is hardly necessary to consider expert opinion.”); Freyer, *supra* note 44, at 443 (“The ICC reminds us that expert evidence is not a requirement of an arbitration, and should only be introduced to the extent necessary to inform the tribunal on a key issue in dispute.”).

183. See *supra* notes 87 & 93 and accompanying text.

tion¹⁸⁴ or on the date of the award?¹⁸⁵ What evidence exists that eating meat from farm animals that were treated with hormones for growth-promotion increases the risk of cancer in humans?¹⁸⁶ What is the maritime boundary between the continental and fisheries zones of two countries?¹⁸⁷ Has construction of man-made islands negatively affected surrounding coral reef systems?¹⁸⁸ Knowing the question to be answered in precise and clear terms provides an essential basis for determining what evidence would be helpful in answering it.¹⁸⁹ This is second nature when evaluating tribunal-appointed experts¹⁹⁰ but no less important when evaluating party-appointed experts.

In formulating potential questions for an expert, decision makers should always acknowledge the extent to which a legal rule has informed the question scope, lest they open themselves up to criticisms that they are unduly allowing expert evidence to influence those rules¹⁹¹ or requiring

184. See, e.g., *Ebrahimi v. Iran*, Award, 30 Iran-U.S. Cl. Trib. Rep. 170, 182–83 (1994); *Starrett Housing Corp. v. Iran*, 16 Iran-U.S. Cl. Trib. ¶ 4.

185. See, e.g., *Siemens A.G. v. Argentine Republic*, ICSID Case No. ARB/02/08, Award, ¶¶ 352–353 (Feb. 6, 2007); *Ger. V. Pol.*, 1928 P.C.I.J. (ser. A) No. 17, at 47, 51–54 (Sept. 13).

186. See *US — Continued Suspension*, *supra* note 111, ¶ 200 (examining whether scientific evidence “focus[ed] on” and “address[ed] . . . the carcinogenic or genotoxic potential of the residues of those hormones found in meat derived from cattle to which the hormones had been administered for growth promotion purposes – as required by paragraph 4 of Annex A of the SPS Agreement”).

187. E.g., *Delimitation of the Boundary in the Gulf of Maine Area*, Judgment, 1984 I.C.J. Rep. 263 (“What is the course of the single maritime boundary that divides the continental shelf and fisheries zones of Canada and the United States of America?”).

188. See *Philippines v. China*, Award, PCA Case No. 2013-19 (2016), ¶ 84 (2015).

189. See *Jacur*, *supra* note 181, at 451 (“The adjudicator should properly identify the extent of the scientific matter at stake and then accordingly carefully pose the questions to the [tribunal-appointed] experts.”); *Pauwelyn*, *supra* note 77, at 186

The first step [in evaluating quantitative economic data] is precisely formulating the relevant question, that is, a question whose answer will be relevant in deciding the applicable legal criterion (e.g., what was the effect of US cotton subsidies provided in year(s) X on world prices of cotton so we can figure out whether these subsidies caused ‘serious prejudice’ in the sense of SCM [the WTO Agreement on Subsidies and Countervailing Measures] Article 6.3(c)).

190. See, e.g., *Corfu Channel (U.K. v. Alb.)*, Judgment, 1949 I.C.J. Rep. 4, 142–69 (Dec. 15) (posing a series of questions to the tribunal-appointed experts); see also *Lukasz Gruszczynski, The Role of Experts in Environmental and Health-Related Trade Disputes in the WTO: Deconstructing Decision-Making Processes*, in *THE ROLE OF “EXPERTS” IN INTERNATIONAL AND EUROPEAN DECISION-MAKING PROCESSES: ADVISORS, DECISION MAKERS OR IRRELEVANT ACTORS* 216, 221–22 (M. Ambrus et al. eds., 2014) (explaining that the process of consulting experts in WTO disputes under the SPS Agreement “starts with the determination of the necessary fields of expertise and is followed by the identification of potential candidates” and the “draft[ing] [of] questions relating to the specific scientific or technical aspects of a dispute”).

191. See *Howse*, *supra* note 32, at 2345–46 (faulting the WTO Panel in *Australia — Salmon* for “never plac[ing] the issues [for expert evaluation] in the context of the legal meaning of the SPS Agreement provisions on risk assessment and scientific evidence” and for “fail[ing] to articulate how it intended to use or weigh the scientific evidence in deciding

experts to answer questions outside their technical competence.¹⁹² When formulating a question for damages experts, for example, a decision maker could, depending on the case, specify the exact contours of the enterprise or other asset to be valued, the relevant valuation date or dates, and whether certain events or facts should be considered or not considered under the applicable legal rules.¹⁹³

Does non-expert evidence or a legal rule obviate the need for expert evidence? With a question in hand, the decision maker should then determine whether available non-expert evidence is sufficient to answer that question or whether the question is purely legal in nature, so that expert evidence is not useful to the decision-making process. In boundary disputes between neighboring countries, for example, decision-makers have often turned to expert evidence only after acknowledging that other sorts of evidence are not enough to draw a definitive boundary.¹⁹⁴ In disputes centering on the design, construction, or operation of industrial facilities or extractive concessions, decision-makers have often determined that expert evidence is required.¹⁹⁵ Similarly, when it comes to calculating the amount of damages owed, decision makers often have found documentary evidence alone insufficient to make an appropriate assessment.¹⁹⁶ Experts

the question of adequacy”); *see also* Appellate Body Report, *Australia — Salmon*, ¶ 6, WTO Doc. WT/DS18/AB/R, (adopted Oct. 20, 1998).

192. *See* Howse, *supra* note 32, at 2347 (“The scientists called upon in *Salmon* were placed in a virtually impossible position: they were asked to make a purely technical/scientific judgment about the adequacy of risk assessment as a *regulatory* tool.”).

193. *See* RIPINSKY & WILLIAMS, *supra* note 43, at 179 (“Tribunals can improve comparability of expert reports by providing guidance on the essential parameters of the task to be performed, for example on: exact object of valuation; relevant valuation date or dates; and whether the impact of certain events subsequent to the valuation date should be accounted for.”).

194. *See, e.g., Honduras Borders (Guat. v. Hond.)*, 2 R.I.A.A. 1307, 1353 (1933) (noting that the Court had commissioned a report from party-appointed experts in light of the inadequacy of existing topographical data); MOORE, *supra* note 25, at 297–98

No geographical map was annexed to the treaty . . . and in our opinion none which can be adopted as a sure and complete proof of the intentions of the negotiators of the treaty. . . . [D]uring the arbitration proceedings no map was produced which was recognized as being entirely accurate by both parties. . . . [The parties] discussed much about the importance and accuracy of their maps, but . . . these discussions did not lead to any decided conclusion as to the value to be given to one of these maps more than to the other as regards the various features of the frontier.

195. *Harza v. Iran*, Interlocutory Award, 2 Iran-U.S. Cl. Trib. Rep. 68, 71 (1983)

[T]he Tribunal cannot decide the liability issues raised by the various alleged defects of site selection, engineers’ drawings, and actual construction of major dams and related irrigation networks without a proper technical description of the alleged defects and of the causes from which they arose. This will require opinions to be sought at least from two experts.

196. In the *Lighthouses Arbitration*, for example, the Permanent Court of Arbitration determined that expert testimony was necessary because “the documents so far lodged by the parties” were not “sufficient to determine” or “capable of furnishing an exact assessment” of

have also often been deemed useful on questions related to the business and commercial practices in a particular industry.¹⁹⁷

By contrast, in a series of cases involving the delimitation of maritime boundaries, the ICJ determined that a legal rule—which generally enables a country to claim a continental shelf up to as far as 200 miles from its coast—meant that it need not weigh expert evidence on the specific geological characteristics offshore of each disputing country.¹⁹⁸ The ICJ instead reasoned that its function is “to make use of geology only so far as required for the application of international law.”¹⁹⁹

Does the expert (or would an expert) have specialized knowledge beyond that of the decision-maker? Evidence offered as “expert” will in general aid in the decision-making process only if the witness has specialized knowledge, skill, or training that complements and goes beyond the competence and training of the adjudicator in answering the specific question presented.²⁰⁰ Two of the first modern cases to use experts extensively in their decision-making processes illustrate the importance of this question. First, in the *Manica Plateau Arbitration*, a boundary dispute, the arbitrator explained that he had sought out an expert “specially qualified in ques-

the total costs and expenditures of the concession at issue. See *Affaire relative à la concession des phares de l'Empire ottoman*, (Fr. v. Greece) 12 R.I.A.A. 155, 228 (Perm. Ct. Arb. 1956) (unofficial translation by author).

197. *Harza v. Iran*, 2 Iran-U.S. Cl. Trib. at 71 (“[I]t appears from the pleadings and the substantiating evidence submitted by both Parties that the reasons for the disagreement between the Parties need to be clarified to the Tribunal by an expert familiar with the business practices and administration of important consulting engineering contracts.”).

198. See, e.g., *Continental Shelf (Libya v. Malta)*, Judgment, 1985 I.C.J. REP. 13, 36 (June 3); *Continental Shelf (Tunis. v. Libya)*, Judgment, 1982 I.C.J. REP. 18, 41, 53, 58 (Feb. 24).

199. See *Tunis. v. Libya*, 1982 I.C.J. at 54; see also Tullio Scovazzi, *Between Law and Science: Some Considerations Inspired by the Whaling in the Antarctic Judgment*, 14 QUESTIONS INT'L L.: ZOOM-IN 13, 21 (2015) (noting that, in these maritime boundary cases, “where it is possible to simply excessively complex questions for the sake of legal clarity, one cannot but agree with the approach taken by the I.C.J. in disregarding scientific discussions and experts’ elaborations – be they appointed by the parties or by the Court”).

200. See, e.g., *Prosecutor v. Perišić*, Case No. IT-04-81-T, Decision on Defense Motion to Exclude Expert Reports of Robert Donia, ¶ 6 (Int’l Crim. Trib. For the Former Yugoslavia Oct. 27, 2008) (noting that the jurisprudence of the International Criminal Tribunal for the former Yugoslavia defines an expert as “a person whom by virtue of some specialised knowledge, skill or training can assist the trier of fact to understand or determine an issue in dispute”); *I.V.E.M. Claim*, 22 Int’l L. Rep. 875 (1955) (noting that, according to the Franco-Italian Conciliation Commission, international adjudication often leads to “inquiries and assessments which presuppose technical knowledge denied to [their own members]”); CHARLES N. BROWER & JASON D. BRUESCHKE, *THE IRAN-UNITED STATES CLAIMS TRIBUNAL* 198 (1998) (“In many international arbitrations, including the Tribunal, the arbitrators are not specialists in the underlying subject matter. Given the broad spectrum of its cases and the complexity of issues presented for adjudication, it therefore not surprising that the use of informed analysis of experts has been invaluable to the Tribunal’s understanding and resolution of . . . issues.” (citations omitted)); Scovazzi, *supra* note 199, at 16 (“Nor can the members of a court, who are experts in law, be supposed to have a universal knowledge, in order to reach by themselves conclusions that require scientific and technical expertise. Here scientific and technical experts have a role to play.”).

tions of geography and topography” because the arguments of the parties were not within his training or competence but were instead “of an essentially technical character.”²⁰¹ As with many boundary disputes, *Manica Plateau* centered on “technical” questions of geography, topography, and hydrology, and how to account for the myriad, overlapping features on the ground—rivers, watersheds, ravines, valleys, slopes, mountains, and plateaus—that defy simple legal analysis.²⁰²

Second, in the *Corfu Channel* case, the ICJ asked naval officers to opine on whether a mine-laying operation in a strait that had damaged a British ship could have been conducted without Albanian authorities on land seeing or hearing the operation, how recently the mines had been laid, and whether the mines were floating or moored.²⁰³ Without prior experience with such operations, it would not be possible to assess the type of operation necessary, its visual footprint, or the level of noise it would create, nor would a layperson have the knowledge or experience to carry out a re-creation and examination of a naval mine-laying operation.²⁰⁴

2. Does the witness have the requisite experience and training in the chosen methodology?

If expert evidence will aid the decision-making process, a decision-maker should next determine whether the expert has the necessary credentials and experience in the relevant field to act as a reliable authority.²⁰⁵ Evaluating the appropriate range of expertise can help to neutralize strategies designed to obfuscate genuine issues, and provide decision-mak-

201. *Manica Plateau Arbitration* (Gr. Brit. v. Port.), 28 R.I.A.A. 283, 308 (1897). The focus of this article is on non-legal experts, although courts and tribunals also often consider testimony on municipal and international law. *But See* Andrew Newcombe, *The Strange Case of Expert Legal Opinions in Investment Treaty Arbitrations*, KLUWER ARBITRATION BLOG (Mar. 18, 2010), <http://kluwarbitrationblog.com/2010/03/18/the-strange-case-of-expert-legal-opinions-in-investment-treaty-arbitrations>.

202. *See* 28 R.I.A.A. 283, 308 (1897); *see also I.V.E.M. Claim*, Final Decision, V RECUEIL DES DÉCISIONS DE LA COMMISSION DE CONCILIATION 153, 177 (Mar. 7, 1955) (explaining that experts were appointed because the issue of valuation was of a financial, economic, and technical nature).

203. *Corfu Channel* (U.K. v. Alb.), Judgment, 1949 I.C.J. Rep. 4, 13, 14, 21 (Dec. 15).

204. *See* Scovazzi, *supra* note 199, at 16

[T]here is no doubt that naval officers [such as those appointed in the *Corfu Channel* case] are better suited than anyone else to clarify whether the operation of laying of mines in a given straight by a third subject can be seen and heard by the look-out posts on the coast. Judges sitting in courts are not likely to engage themselves in activities of investigation on moonless nights, such as the test of visibility carried out by the committee of experts in the Strait of Corfu.

205. NAOMI ORESKES & ERIK M. CONWAY, *MERCHANTS OF DOUBT: HOW A HANDFUL OF SCIENTISTS OBSCURED THE TRUTH ON ISSUES FROM TOBACCO SMOKE TO GLOBAL WARMING* 272 (2010) (“[W]e need to pay attention to who the experts actually are – by asking questions about their credentials, their past and current research, the venues in which they are subjecting their claims to scrutiny, and the sources of financial support they are receiving.”).

ers with the opportunity to appoint additional experts or to require the parties to disclose the limits inherent in their choice of experts.

To supply testimony that can be considered reliable, an expert witness must have more than general knowledge of the topic on which they are opining.²⁰⁶ They must have experience in the particular type of problem at issue in a dispute,²⁰⁷ partly because so many tasks in the real world do not “correspond to the speciality of any single profession.”²⁰⁸ Experience in advising on corporate acquisitions of oil field properties, for example, would not mean that a witness is qualified to testify on other aspects of the oil and gas industry, such as the complex hedging of risk through futures trading.²⁰⁹ Nor is an appraiser “at home in the whole field of appraisal[:] he is a real-estate appraiser, a mine appraiser, a security analyst.”²¹⁰

As for what counts as experience in a reliability analysis, what matters is an expert’s previous experience with the chosen methodology outside the dispute context. At the same, an expert need not have conducted or supervised the specific studies on which her testimony relies.²¹¹ If the question is whether certain human activities can cause damage to coral reefs, for example, an expert is more likely to provide reliable answers to that question if she has experience in evaluating anthropogenic effects on coral reefs as part of an academic or other study, although she need not have studied the particular coral reef at issue in the dispute.²¹² An expert asked to opine on the considerations that go into construction of large dams is more likely to provide reliable answers if she has real-world expe-

206. Prosecutor v. Perišić, Case No. IT-04-81-T, Decision on Defense Motion to Exclude Expert Reports of Robert Donia, ¶ 8 (Int’l Crim. Trib. For the Former Yugoslavia Oct. 27, 2008) (“[S]tatements or reports of an expert witness will only be treated as expert evidence, insofar as they are based on the expert’s specialised knowledge, skills or training. Statements that fall outside the area of expertise will be treated as personal opinions of the witness and will be weighted accordingly.”); see also ORESKES & CONWAY, MERCHANTS OF DOUBT, *supra* note 205, at 271 (“An all-purpose expert is an oxymoron.”).

207. See WHITE, *supra* note 9, at 148 (noting that, “[i]n the field of valuation of property, the qualified expert is likely to be a specialist in a particular type of valuation problem”).

208. See JAMES C. BONBRIGHT, 1 THE VALUATION OF PROPERTY 7 (1937) (considering the task of asset valuation).

209. See Proton Energy Group SA v. Orlen Lietuva, (2013) EWHC (Comm) 2872, QB (Eng.).

210. See BONBRIGHT, *supra* note 208, at 7.

211. The WTO Appellate Body has come to the opposite conclusion in at least one case. See *US — Continued Suspension*, *supra* note 111, ¶ 199 (discounting an expert’s testimony in part because the underlying scientific studies were not “carried out by him under his supervision”).

212. See *Phil. v. China*, Case No. 2013-19, Award, ¶ 821 (Perm. Ct. Arb. 2016) (noting that a tribunal-appointed expert asked to opine on the potential effects of China’s construction of man-made islands on surrounding coral reefs was a “coral reef ecologist with over ten years’ research experience in Southeast Asia [and] the Pacific Islands . . . and ecological work . . . focused on coral reef restoration and ecological functioning and the impact of environmental and anthropogenic factors on coral reef benthic communities”).

rience in the construction of large dams, although she need not have participated in the construction of the dam at issue in the dispute.²¹³

Any assessment of an expert's reliability must also include a thorough examination of any previous participation as an expert in other legal disputes.²¹⁴ The fact that the expert's previous participation in other disputes was consistent with the evidence that she provides in the dispute at hand and that other legal decision makers relied on an expert's use of the same or similar methodology can be an indicator of reliability, although no substitute for a decision maker's own reliability analysis.

Confidentiality associated with some forms of international dispute settlement, including commercial arbitration, may make it difficult to assess reliability on the basis of an expert's prior participation in international disputes. But agreements and procedural rules that increasingly emphasize the benefits of disclosing decisions, briefs, and witness reports and statements may enable decision makers to more easily track and consider the history of experts.²¹⁵ And a decision maker can also of course look to an expert's participation in domestic court proceedings or domestic regulatory rulemaking.

When the parties do not agree on the sort of expertise required, a decision maker may decide to appoint (in the case of tribunal-appointed experts) or weigh (in the case of party-appointed experts) a wider range of experts who reflect those disparate views.²¹⁶ But a decision maker should remain aware that the parties' views may reflect strategic choices intended

213. See *Harza v. Iran*, Interlocutory Award, 2 Iran-U.S. Cl. Trib. Rep. 68, 71 (1983) (determining, in a case concerning the site-selection for and construction of major dams and related irrigation networks "will require opinions to be sought at least from two experts, one experienced in the field of geology and rock mechanics and the other in hydraulics, *both of whom are also experienced in the building of major dams*" (emphasis added)).

214. See, e.g., *Prosecutor v. Perišić*, Case No. IT-04-81-T, Decision on Defense Motion to Exclude Expert Reports of Robert Donia, ¶ 12 (Int'l Crim. Trib. For the Former Yugoslavia Oct. 27, 2008).

215. For example, mechanisms such as the United Nations Convention on Transparency in Treaty-based Investor-State Arbitration, through application of the UNCITRAL Rules on Transparency to a potentially large set of investment arbitrations, may increase the database of expert reports and international decisions relying on expert evidence. See G.A. Res. 69/116 (Dec. 10, 2014) (applying the UNCITRAL Rules on Transparency to investor-state disputes where both the respondent and the claimant's state are parties to the Convention, and the treaty was concluded before April 1, 2014); G.A. Res. 68/109 (Dec. 16, 2013) (providing that "expert reports and witness statements, exclusive of the exhibits thereto, shall be made available to the public, upon request by any person to the arbitral tribunal").

216. See Panel Report, *EC — Asbestos*, ¶ 5.2 WTO Doc. WT/DS135/R and Add. 1 (adopted 5 April 2001) (describing, in a dispute over whether a French measure banning the marketing of asbestos products were inconsistent with WTO rules, the Canadian preference for experts specializing in the comparative toxicity of asbestos and non-asbestos fibers, and the EC preference for cancer epidemiologists, including specialists in asbestos-caused cancers); *US — Continued Suspension*, *supra* note 111, ¶ 7.91 (describing the EC's view that the panel should have appointed an expert in animal science, including veterinary practices on administration of growth hormones).

to influence the outcome of a dispute,²¹⁷ and may not encompass the full range of expertise necessary to evaluate the questions presented.

3. What methodologies does the relevant technical field consider reliable to answer the question presented?

Once a question has been formulated and a decision maker has ascertained itself of an expert's qualifications, the next step is to determine whether the expert's *methodology* is considered reliable in the relevant technical field. Most technical fields rely on "peer review"—evaluation by other specialists in the field—to determine whether work is potentially valid.²¹⁸ The fact that a methodology has been subject to and withstood the scrutiny of peer review can signal that other experts in the relevant field have analyzed the methodology, determined that the methodology is capable of reliable application, and generally view the methodology as credible.²¹⁹ At base, peer review determines "[w]hat counts as knowledge" in technical and specialized fields.²²⁰ The goal of peer review is to assess the quality of analysis, including the design and execution of any studies on which that analysis is based,²²¹ and to identify gaps and limits in knowledge and the need for further investigation.²²²

Because "peer review provides a benchmark of legitimacy," it can provide a basis for assessing the reliability of expert evidence in international disputes.²²³ As the need to rigorously examine the reliability of expert

217. See Gruszczynski, *supra* note 190, at 228 ("[A] preference for a specific type of expertise may reflect deliberate and strategic choices of the parties aimed at influencing the ultimate outcome of a dispute.").

218. See Orellana, *supra* note 180, at 65 ("Questions of scientific truth are the function of peer review, which serves as a symbol of professional accountability and ensures the democratic control of science.").

219. See Orellana, *supra* note 180, 57 ("[A]lthough the political nature of the Intergovernmental Panel on Climate Change has been evident since its creation, the institutional mechanisms designed for discussion between scientists and policy makers, coupled with extended peer review, have significantly contributed to the panel's credibility.").

220. See ORESKES & CONWAY, *MERCHANTS OF DOUBT*, *supra* note 205, at 269 (noting that peer review determines what scientists consider to be scientific knowledge).

221. See FOSTER, *supra* note 3, at 10.

222. See Wirth, *supra* note 133, at 842

[T]he scientific peer review process operating in a regulatory context can reduce disagreement, identify gaps and holes, and articulate the need for further investigation. . . . [P]eer review is responsive to a characterization of science as an ongoing search for knowledge against a constantly shifting and evolving background that by its very nature is always operating at new frontiers.

223. See Orellana, *supra* note 180, at 65 ("[T]ribunals facing the difficult task of recognizing the scientific character of evidence will find that peer review provides a benchmark of legitimacy that justifies qualified deference."); Jacur, *supra* note 181, at 452 ("Peer review is, in my view, a reliable instrument to ensure the credibility of experts."); MARBOE, *supra* note 64, at 183 (explaining that, in investment arbitrations, "an arbitral tribunal needs basic knowledge about the most important valuation principles and methods which are known and applied in international valuation practice"); cf. Céline Lévesque, *Science in the Hands of International Investment Tribunals: A Case for "Scientific Due Process"*, in 20 FINNISH Y.B.

evidence has increased with the complexity of international disputes over the years, legal decision makers have begun to rely more heavily on the relevant technical communities for guidance.²²⁴ This has mirrored a similar trend in the domestic law context.²²⁵ Parties to international agreements now hardwire reliance on technical communities into the legal text itself.²²⁶

But peer review is far from perfect,²²⁷ and a legal decision maker should insist that each expert demonstrate that the testimony is supported by peer-reviewed primary materials and that those materials as a whole meet minimum acceptable standards for reliability in the field,²²⁸ which will generally mean that the research was conducted in a disinterested manner, shared with other researchers, and subjected to disinterested evaluation by peers.²²⁹ Decision makers should analyze the reliability of

INT'L L. 259, 288 (2011) (“[A] tribunal might consider whether . . . regulatory science was peer reviewed.”).

224. See e.g., *Philip Morris Brands Sàrl v. Oriental Republic of Uruguay*, ICSID Case No. ARB/10/7, Award, ¶ 392 (2016) (noting, as detailed by Uruguay’s experts, that evidence of consumers’ misconceptions of health effects of “light” and “low-tar” cigarettes available at the time Uruguay adopted the challenged measures included “[n]umerous scientific studies . . . published . . . in leading international journals” and “cited by the U.S. Surgeon General and the U.S. National Cancer Institute” (citations omitted)); *Whaling in the Antarctic (Austl. v. Japan: N.Z. intervening)*, Judgment, 2014 I.C.J. Rep. 226 (Mar. 31) (separate opinion of Bhandari, J.) (recognizing peer review as an indicator of “reliable science”); *Methanex Corp. v. United States*, Final Award of the Tribunal on Jurisdiction and Merits, ¶ 101 (NAFTA Ch. 11 Arb. Trib. 2005), <http://www.italaw.com/sites/default/files/case-documents/ita0529.pdf> (noting that an environmental impact report had been subject to “public hearings, testimony, and peer review,” “emerging as a serious scientific work from such an open and informed debate”).

225. See THOMAS O. MCGARITY & WENDY E. WAGNER, *BENDING SCIENCE: HOW SPECIAL INTERESTS CORRUPT PUBLIC HEALTH RESEARCH* 1 (1st ed. 2008) (“[L]egal decision makers have struggled to develop more rigorous tools for assessing the reliability of the scientific information that informs health policies. The solution, most have agreed, is for decision makers to rely more heavily on the scientific community for oversight and assistance.”).

226. Foster, *supra* note 45, at 151 (reasoning that the “risk assessment” provision in the SPS Agreement requires a WTO dispute settlement panel to “assess whether the science relied on by a Member is legitimate science according to the standards of the relevant scientific community and also scrutinize the reasoning in a risk assessment with reference to the underlying science”).

227. See generally John P.A. Ioannidis, *Why Most Published Research Findings Are False*, 2 PLoS MED. 696, 700–01 (2005); William A. Wilson, *Scientific Regress*, FIRST THINGS (2016), <https://www.firstthings.com/article/2016/05/scientific-regress> (observing that “[a]t its best, science is a human enterprise with a superhuman aim” but noting “example after example of how the human element of this enterprise harms and damages its progress, through incompetence, fraud, selfishness, prejudice, or the simple combination of an honest oversight or slip with plain bad luck”).

228. See Ioannidis, *supra* note 227, at 700-01.

229. See MCGARITY & WAGNER, *supra* note 227, at 47 (setting out these “informal norms of good scientific practice”); see also Markus Wagner, *Law Talk v. Science Talk: The Languages of Law and Science in WTO Proceedings*, 35 FORDHAM INT’L L.J. 151, 157 (2011) (“[S]cientific inquiry is disinterested in the policy outcome that follows any revelation, through publication or otherwise, of the results of the scientific inquiry.”).

underlying evidence in view of the totality of the evidence presented, without focusing on the outcomes of particular studies.²³⁰

How can non-experts discern when peer-reviewed studies are reliable? A widely-cited analysis of peer-reviewed scientific research concluded that research findings are more likely to be true when study sample sizes are larger (e.g., the relatively large randomized controlled trials in cardiology) and the demonstrated effects of studied phenomenon are larger (e.g., the relatively large effects of smoking on cancer or cardiovascular disease).²³¹ Similarly, research findings in a particular area are more likely to be true when confirmatory studies have been performed (e.g., meta-analyses of previous studies) and adhere to standardized conduct and reporting practices (e.g., those that decrease the discretion and flexibility of individual researchers in designing and reporting studies).²³²

Certain sorts of evidence demand heightened scrutiny, particularly those where financial and other interests may unduly influence the research process. For instance, when experts rely on “policy-relevant” research, technical knowledge that may influence government policy or litigation outcomes,²³³ they should demonstrate that the relevance of the outcomes did not compromise the research process.

An advocate of a particular policy outcome can direct the outcome of policy-relevant research by commissioning studies and then exerting significant control so that the studies work backward from the advocate’s desired outcome.²³⁴ Or the advocate may simply re-package existing studies in a light more favorable to the advocate, including by cherry-picking

230. See Ioannidis, *supra* note 227, at 700–01 (“[M]ost research questions are addressed by many teams, and it is misleading to emphasize the statistically significant findings of any single team. What matters is the totality of the evidence.”); Christie Aschwanden, *Science Isn’t Broken: It’s Just a Hell of a Lot Harder than We Give It Credit For*, FIVETHIRTYEIGHT (Aug. 19, 2015), <https://fivethirtyeight.com/features/science-isnt-broken/> (“The important lesson here is that a single analysis is not sufficient to find a definitive answer.”).

231. Ioannidis, *supra* note 227, at 697–98.

232. See *id.*; see also Aschwanden, *supra* note 230 (describing the “decisions scientists make as they conduct a study” that can unduly influence research outcomes).

233. See MCGARITY & WAGNER, *supra* note 225, at 10–11 (2008); see also Bonneuil & Levidow, *supra* note 49, at 77 (“[N]either regulatory science nor research science can claim to stand free from historical, political and cultural situations.”); Ioannidis, *supra* note 227, at 698 (“Conflicts of interest and prejudice may increase bias. Conflicts of interest are very common in biomedical research, and typically they are inadequately and sparsely reported. . . . The hotter a scientific field . . . the less likely the research finding are to be true.”).

234. See MCGARITY & WAGNER, *supra* note 225, at 8–9, 38–39 (describing this phenomenon in the context of scientific research); Werner, *supra* note 33, at 48

The distrust in expert input can partly be explained by the fact that scientific experts sometimes ally with specific commercial firms, interest[] groups and political parties. One of the most disturbing examples are the so-called ‘merchants of doubt,’ a group of scientific experts with strong links to industry and conservative politicians who proved willing to cast doubt on results of scientific research in the areas of smoking tobacco, acid rain, the depletion of the ozone layer and global warming.

research.²³⁵ Because policy-relevant research often addresses narrow questions arising from policy debates, advocates—especially regulated industries—may be the only ones interested in pursuing or reviewing it, or to devote sufficient resources to it.²³⁶ This may lead to asymmetrical representation of views in those areas, in favor of the outside advocates' position, as well as little disinterested examination of the details of the research.²³⁷ Outside advocates know that funding policy-relevant research can then receive the imprimatur of legitimacy through peer review because lower-tier publications are eager for submissions and may subject research to only cursory review.²³⁸

If it appears that research supporting expert evidence is “policy-relevant,” the next question is whether outside advocates—with an interest in the outcome—funded or otherwise participated in the research or peer-review process. Although the mere fact that advocates of a particular outcome have funded policy-relevant research does not mean that it is unreliable,²³⁹ in many cases advocate funding does influence research outcomes.²⁴⁰ For that reason, advocate funding, especially when coupled with contractual control over research results or publication, or an advocate's significant collaboration in the research process, require a legal decision maker to look further into the provenance of empirical evidence. At a minimum, generally accepted norms require disclosure of research sponsorships or other potential financial interests in the subject matter, and confirmation that the named authors of a publication directed and conducted the underlying research.²⁴¹ A tribunal or court should allow the

235. See MCGARITY & WAGNER, *supra* note 225, at 39.

236. See *id.* at 11–13, 54–55.

237. See *id.* at 54–55 (“[S]tudies that sponsors control in obvious ways may attract little attention from independent scientists because they are reluctant to spend their valuable time refuting studies that most of their peers know to be unreliable.”).

238. See *id.* at 53 (noting that “many second- and third-tier [scientific] journals are hungry for submissions and will publish articles that receive minimal peer scrutiny both before and after publication.” (citation omitted)).

239. See *id.* at 65

Independent scientists can and regularly do conduct policy-relevant research in a disinterested way, without input from affected parties or financial inducements that cause them to tilt or skew the research toward a particular end. Furthermore, scientists can and do carry out sponsored research in accordance with the norms and procedures of science without being influenced by the sponsors' economic interests.

(citation omitted).

240. See *id.* at 95–96 (detailing the “funding effect” in biomedical research, which according to a “comprehensive review article summarizing 1,140 biomedical research studies” means that “ ‘industry-sponsored studies were significantly more likely to reach conclusions that were favorable to the sponsor than were non[-]industry studies.’ ” (citation omitted)).

241. See *id.* at 76–77:

Most scientists would agree that when a sponsor contractually controls research or otherwise acts as a significant collaborator, the norms of science require that any

parties to present evidence on the extent to which advocates have commissioned research underlying expert evidence and the degree to which advocates controlled the research outcomes.²⁴² If a key publication of policy-relevant research does not disclose this information, the party relying on the publication should provide those disclosures and explain why any advocate's financial interest or other involvement did not lead to unreliable results.²⁴³

Lastly, although peer review can be a significant indicator of reliability, not all reliable work is peer-reviewed or even published.²⁴⁴ For specialized fields with significant industrial or commercial applications, for example, standards from the non-academic private sector may not be peer-reviewed but will often play an important role in defining whether a particular methodology is generally accepted.²⁴⁵

published version of the study disclose at least the fact of sponsorship The most insidious technique for obscuring provenance is the practice commonly employed in the pharmaceutical industry of hiring ghost-writing companies that in turn hire bright young science writers to turn data and analyses from company-sponsored clinical trials into articles suitable for publication in medical journals under the signatures of prominent researchers.

242. *See id.* at 240

The [U.S.] courts should allow the parties to conduct discovery and present evidence on the extent to which some or all of the research underlying an expert's testimony was in fact commissioned by an entity with a direct or indirect interest in the litigation and on the degree to which that entity exercised control over the outcome of the research.

243. *See id.* at 239 (“[C]ourts and agencies should at the very least ensure that the provenance of all research that they rely on to support regulatory and judicial determinations is, as far as is possible, divulged and considered in the decision-making process.”); Biases and errors may stem from sources other than advocate funding. Specialists in a particular field may have biases due to their commitments to certain theories or to certain scientific organizations. Legal decision makers should take into account whether any potential biases from these other sources have undermined the reliability of the expert's conclusions. Ioannidis, *supra* note 227, at 698 (“Prejudice may not necessarily have financial roots. Scientists in a given field may be prejudiced purely because of their belief in a scientific theory or commitment to their own findings.”); Theofanis Christoforou, *Settlement of Science-Based Trade Disputes in the WTO: A Critical Review of the Developing Case Law in the Face of Scientific Uncertainty*, 8 N.Y.U. ENVTL. L.J. 622, 630 (2000) (“[S]cientists coming from [specific] organizations may be unfairly biased in favor of maintaining their organization's standards and recommendations”); *see also* Gruszczynski, *supra* note 32, at 227 (“[P]revious affiliation should be regarded as an important element when assessing the existence of potential conflicts of interest.”).

244. *See* Matthew W. Swinehart, *Remedying Daubert's Inadequacy in Evaluating the Admissibility of Scientific Models Used in Environmental-Tort Litigation*, 86 TEX. L. REV. 1281, 1307 (2008) (cautioning against similar misuse of peer review as a factor in analyzing expert reliability).

245. FOSTER, *supra* note 3, at 9 (“Standards emanating from the private sector may also have an effect within international law.”).

4. Which expert's methodology is most reliable in this context?

Once a tribunal has satisfied itself that the methodology is one that is generally accepted and used in the field to address the sort of question at issue, a reliability analysis should then ask whether the expert has chosen a methodology that appropriately fits the particular circumstances of the case and has explained why that methodology is preferred over other available or proffered methodologies.²⁴⁶

Views of the Relevant Technical Field—To use expert evidence to make decisions, decision-makers need not arrive at a unified theory by synthesizing or reconciling competing expert evidence, but must assess the relative reliability of each expert submission.²⁴⁷ This is critical given that an evaluation of a methodology's reliability will rarely if ever lend itself to a binary analysis, resulting in a definitive determination that particular methodologies are "reliable," while others are "not reliable." Instead, the reliability of expert methodologies is a matter of degree, and the degree of each methodology must be compared to that of others.²⁴⁸ To that end, a decision maker should identify and examine the generally accepted norms and standards of the expert's field,²⁴⁹ including the relative reliabilities of

246. See Pauwelyn, *supra* note 77, at 186 ("[T]he pros and cons [of empirical economic methodologies] should be made explicit, including potential statistical and identification problems" and "[a]lternative methodologies should be discussed."); RIPINSKY & WILLIAMS, *supra* note 43, at 192 ("There is no single valuation method that fits all cases and scenarios. This gives tribunals flexibility but poses the problem of selecting a method that is appropriate and justified in a particular case.")

247. See, e.g., Panel Report, *US — COOL*, ¶ 7.513, WTO Doc WT/DS384, 386/R (Nov. 18, 2011) (noting, in evaluating competing economic studies, that it is not the task of a WTO panel, "to establish a unified econometric report or to conduct our own econometric assessment; instead we will assess the robustness of each study"); *CMS Gas Transmission Company v. Argentine Republic*, ICSID Case No. ARB/01/8, Award, ¶¶ 411–417 (May 12, 2005) (explaining why the tribunal chose to value the business at issue using the DCF method rather than other approaches such as the multiples, comparables, option-value, and asset-value methods); *Starrett Housing Corp. v. Iran*, Final Award, 16 Iran-U.S. Cl. Trib. Rep. ¶¶ 282, 336 (1987) (noting approvingly that an expert employed "generally accepted accounting principles" and "generally recognized valuation practices").

248. See Pauwelyn, *supra* note 77, at 26 ("[T]he scope, relevance, assumptions and limits of each study should be examined and, on that basis, each should be given appropriate weight (or, as the case may be, no weight at all)."); cf. CHRISTOPH H. SCHREUER ET AL., *THE ICSID CONVENTION: A COMMENTARY* 1012 (2d ed. 2009) ("The speculative character of damages theories in the calculation of lost profits is a matter of degree.")

249. See, e.g., *EC—Asbestos* *supra* note 216, ¶ 5.147 (determining that it was generally accepted among scientists that there was "a direct and linear relationship between the relative risk of lung cancer and cumulative exposure to asbestos"); *CMS Gas Transmission Company*, ICSID Case No. ARB/01/8, Award, ¶ 416 (adopting the DCF valuation method in part because it is "universally accepted"); *Harza v. Iran*, 2 Iran-U.S. Cl. Trib. Rep. at 76, ¶ 99 (1986) (declining to rely on an expert opinion that the claimant could have improved its geological assessment in preparing for the construction of a dam because "the evidence indicates, no clear technical or professional standards exist" as to the necessary elements of such an assessment in that context). The idea that tribunals should consider the "most common and accepted methods" in making damages assessments comes from the legal standard of "fair market value." See RIPINSKY & WILLIAMS, *supra* note 43, at 211 ("If one really wishes to estimate the *market* value of an investment, then one should use most common and ac-

different methodologies as gleaned from peer-reviewed studies and collected in national and international guidelines.²⁵⁰ A consensus reached among decision-makers in international disputes on the use of a methodology may also indicate that a methodology is the most reliable in certain sorts of cases and fits within the framework of applicable legal rules,²⁵¹ although it is not a substitute for full reliability evaluation.²⁵²

The use of a particular methodology can be adjudged “correct” only after sufficiently widespread acquiescence in the field over time, a “consensus” that is always subject to revision. Many specialized fields by their very nature do not generally offer definitive answers but reflect some degree of uncertainty, due to imperfect knowledge or ever-developing opinions,²⁵³ and center instead on continual evaluation and testing of hypotheses that are understood to be valid only until they are disproved.²⁵⁴ The relevant field may remain in a state of disequilibrium on the acceptance of a methodology, with specialists holding a range of

cepted methods to reach that value, and the DCF method *is* an appropriate method.”); Simons, *supra* note 64, at 234 (“Because tribunals are seeking to determine market value, their decisions should be informed by the real-world practices of willing buyers and willing sellers.”); *id.* at 235 (arguing that tribunals should “abandon the practice of dismissing the DCF method simply because an entity being valued is not a ‘going concern’” in part because “[f]inancial analysts employ the method even for enterprise that arbitral tribunals would not deem ‘going concerns’”). But even absent the legal standard the general acceptance of particular valuation methodologies has also informed broader considerations of the reliability of valuation evidence. See, e.g., Manuel A. Abdala, *Key Damage Compensation Issues in Oil and Gas International Arbitration Cases*, 24 AM. U. INT’L L. REV. 539, 548–49 (2009) (noting that DCF is “the most common methodology used in valuation analyses” in part because “it is widely supported by the professional literature” and “most investors rely on a DCF analysis to determine whether or not to undertake a particular project” (citations omitted)); Starrett Housing, 16 Iran-U.S. Cl. Trib. Rep. at 241 (Holtzmann, J., concurring opinion) (reasoning that “experience in the financial community throughout the world attests to the suitability of the DCF Method in valuing all kinds of businesses regardless of their purposes or the length of time they are expected to operate”).

250. See RIPINSKY & WILLIAMS, *supra* note 43, at 192 (noting that, although “[t]here is no comprehensive and authoritative manual of universal application that tribunals could use to guide them through valuation. . . . there are a number of well-respected academic manuals as well as national and international guidelines on valuation that can be referred to.” (citation omitted)).

251. Cf. CMS Gas Transmission Company, ICSID Case No. ARB/01/8, Award, ¶ 416 (adopting the DCF valuation method in part because it has been adopted by “numerous tribunals”); see also RIPINSKY & WILLIAMS, *supra* note 43, at 192 (“[T]he growing body of arbitral practice need[s] to be reviewed and analysed in order to understand the content and applicability of different valuation methods for the purposes of damages awards in investment arbitrations.”).

252. See *supra* notes 138–39 regarding the discount rate discussion in *Venezuela Holdings*, and accompanying text.

253. See Miles, *supra* note 33, at 157 (“[S]cience is not always able to provide exact, conclusive answers, particularly where there is uncertainty, imperfect knowledge or conflicting scientific opinion.”).

254. See *id.* (noting that “scientific knowledge” is “an organic process of evolving understandings and continuous interactive exchanges of hypothesis and refutation”); Douglas Crawford-Brown et al., *Environmental Risk, Precaution, and Scientific Rationality in the Context of WTO/NAFTA Trade Rules*, 24 RISK ANALYSIS 461, 462 (2004) (“An important aspect

views.²⁵⁵ Consensus may simply remain elusive on a time-scale useful to legal decision makers, who must make decisions today, especially when it comes to cutting-edge scientific research.²⁵⁶

For that reason, expert evidence need not reflect “mainstream” or consensus opinion to be reliable. The *methodology*—not the *conclusion*—should be generally accepted, even if an expert’s conclusions based on that methodology diverge from the opinions of most other experts in the field.²⁵⁷ The WTO Appellate Body has reached a similar conclusion, in a slightly different but analogous context: how to determine when a government has “based” a health measure on a “risk assessment,” consistent with Article 5.1 of the SPS Agreement. The Appellate Body has consistently explained that a government may base a risk assessment not only on mainstream science but also on scientific opinions outside the mainstream.²⁵⁸ At the same time, it has required minority opinions to “have the necessary scientific and methodological rigour to be considered reputable sci-

of science is that it is by nature progressive, with methods designed to approach a better understanding over time.”).

255. See Naoto Jinji, *An Economic Theory of the SPS Agreement*, in SCIENCE AND TECHNOLOGY IN INTERNATIONAL ECONOMIC LAW: BALANCING COMPETING INTERESTS 53, 54 (B. Mercurio & Kuei-Jung Ni eds., 2014) (“[T]he scientific community may fail to reach a consensus [in risk analysis regarding health and safety], as a number of opposing but equally plausible views may exist.”); MCGARITY & WAGNER, *supra* note 225, at 56 (“[I]n some areas of public health and environmental research, scientists are in no position to validate the accuracy of a particular study’s methodological approach at the time its findings are announced and pressed on agencies or courts.”).

256. See Orellana, *supra* note 180, at 55

While scientific method and peer review lead towards consensus as science’s operational criteria, consensus, instead, becomes more elusive the farther out from core knowledge and the farther into the frontier areas of research and experimentation one travels. Such frontier areas are characterized by persistent uncertainty and diverging interpretations, where emerging consensus or minority views may depend on socio-cultural factors and on informal negotiations among scientists.

(citations omitted).

257. See Pauwelyn, *supra* note 48, at 252–54 (reasoning, in the context of WTO disputes over health policy, that there is “no need to find that at least a majority of the scientific community is in favour of a proposed health measure,” even though “the minority opinion must nonetheless come from ‘qualified and respected sources’”); Caroline E. Foster, *The “Real Dispute” in the Southern Bluefin Tuna Case: A Scientific Dispute?*, 16 INT’L J. MARINE & COASTAL L. 571, 589 (2001) (“Developments in multilateral international trade law also recognise that the key features of what may constitute genuine ‘science’ do not necessarily include a requirement that a scientist belong to a ‘mainstream’ scientific community, depending on the circumstances. There is an increased emphasis on scientific principles and methodology.”).

258. See, e.g., *US — Continued Suspension*, *supra* note 111, ¶ 591; see also Lukasz Gruszczynski, *Science and the Settlement of Trade Disputes in the World Trade Organization*, in SCIENCE AND TECHNOLOGY IN INTERNATIONAL ECONOMIC LAW 11, 13 (Bryan Mercurio & Kuei-Jung Ni eds., 2014) (“The SPS case law is consistent in holding that risk assessment may be based not only on mainstream science, but also on the opinions of scientists taking a divergent view.”).

ence.”²⁵⁹ In other words, as one commentator has noted, consideration of minority opinions should involve an “evaluation of a specific claim against more general theories or established scientific views, the logical coherence of a claim, its methodological soundness, or its ability to explain a particular phenomenon.”²⁶⁰

Failure to evaluate in a transparent manner the relative reliability of competing expert evidence can leave a decision vulnerable to criticism. In *Philip Morris v. Uruguay*, tobacco companies challenged Uruguay’s cigarette marketing restrictions, including “single presentation” requirements that prevented the companies from selling one brand (e.g., Marlboro) under more than one variant (e.g., “Light” and “Blue”).²⁶¹ Both sides introduced expert evidence on the tobacco companies’ claim that the requirement was overbroad because, while descriptive brand variants such as “Low Tar,” “Light,” or “Mild” might lead consumers to believe incorrectly that some cigarettes are less harmful than others, other brand variants—such as “Red,” “Gold,” and “Blue”—were not similarly misleading.²⁶²

Two arbitrators rejected the companies’ challenge to the single presentation requirement, partly relying on Uruguay’s expert evidence and an amicus submission from the World Health Organization (WHO) that asserted that brand variants, including different colors, could mislead consumers.²⁶³ The third arbitrator, concurring in part and dissenting in part, found the competing expert evidence, from the tobacco companies, more compelling in key respects.²⁶⁴ None of the arbitrators, however, explained why it viewed one set of expert evidence as more reliable than the other. In concluding that the measure was “not disproportionate” to the public health concern,²⁶⁵ the majority offered in passing that the “rationale” for the requirements was “supported by the evidence” “as held by the WHO”²⁶⁶ and that Uruguay’s experts had “indicated” that the tobacco companies’ promotion of “light” cigarettes as a way to reduce smoking risk “misrepresented what would actually happen.”²⁶⁷

The third arbitrator concurred in the majority’s judgment that Uruguay’s prohibition of descriptive and color brand variants was not disproportionate, finding Uruguay’s evidence on color variants “sufficient,”

259. See *US — Continued Suspension*, *supra* note 111, ¶ 591.

260. See Gruszczynski, *supra* note 258, at 13.

261. See *Philip Morris Brands Sàrl v. Oriental Republic of Uruguay*, ICSID Case No. ARB/10/7, Award, ¶¶ 357–367 (2016).

262. *Id.* ¶¶ 391–392, 406; see also *id.* ¶ 160 (Born, J., concurring in part and dissenting in part).

263. See *id.* ¶ 407 (determining that, “as held by the WHO, ‘the rationale for [Uruguay’s] action [was] supported by the evidence’” (quoting WHO submission)).

264. See *id.* ¶¶ 159–164 (Born, J., concurring in part and dissenting in part).

265. *Id.* ¶ 409.

266. *Id.* ¶ 407.

267. *Id.* ¶ 403.

albeit “very limited,” unimpressive,” and “tenuous.”²⁶⁸ The concurrence did not discuss the content of any of the expert reports or provide substantive reasons for these characterizations except to say that they were based on “the reasons detailed in the Claimants’ expert evidence.”²⁶⁹

The third arbitrator dissented with respect to all other brand variants, such as words or numbers in addition to the brand name, seasonal or geographic variations, or different languages.²⁷⁰ But again the dissenting portions of the opinion disclose nothing about the substance of the expert reports, concluding summarily that “the fact that some uses of colors in some brands of tobacco products may be regarded as misleading in some circumstances does not suggest, even indirectly, that all other variations . . . are also misleading.”²⁷¹ This analysis failed to address the WHO’s assertions that any “brand extension” or “brand variant” can create misleading consumer perceptions, based on peer-reviewed studies purportedly showing that “people try to find attributes among brand variants” and that any feature of pack design can create misconceptions.²⁷² From the face of the arbitrators’ opinions, then, the parties and the public cannot determine whether the arbitrators sufficiently evaluated the expert evidence or the extent to which the evidence factored into their conclusions.

Tribunals in investments disputes have also often failed to evaluate the relative reliability of competing valuation methodologies, particularly when the investment has little or no track record and is thus not considered a “going concern.”²⁷³ The tribunal in *BG Group v. Argentina*, for example, dismissed use of the DCF method as leading “to a result which is uncertain and speculative,” without addressing the details of the expert’s DCF methodology or considering whether specialists in the field would have relied on that methodology in valuing the investment.²⁷⁴ Rather, the tribunal rejected the reliability of the DCF methodology based on two available transactions that the tribunal considered comparable to the investment.²⁷⁵ Tribunals retain broad discretion in determining the standard

268. *Id.* ¶ 161.

269. *Id.* ¶¶ 160–161.

270. *Id.* ¶¶ 162–163.

271. *Id.* ¶ 163.

272. *See Philip Morris Brands Sàrl v. Oriental Republic of Uruguay*, ICSID Case No. ARB/10/7, Written Submission by the World Health Organization and the WHO Framework Convention on Tobacco Control Secretariat, ¶¶ 77–81 (Jan. 28, 2015).

273. *See Simmons*, *supra* note 64, at 226–32 (outlining four decisions between 2007 and 2009 that rejected the DCF method on the ground that it was too speculative given available evidence).

274. *BG Grp. Plc. v. Republic of Argentina*, Final Award, ¶ 439 (UNCITRAL Arb. Proc. 2008), <http://www.italaw.com/sites/default/files/case-documents/ita0081.pdf>.

275. *See id.* ¶¶ 440–444; *see also Siag v. Arab Republic of Egypt*, ICSID Case No. ARB/05/15, Award, ¶¶ 566–570 (May 11, 2009) (rejecting the DCF method in favor of the comparable transaction method without comparing relative reliability of the methods).

for evidence underlying damages calculations,²⁷⁶ but a more robust and defensible analysis would have considered the reliability of the DCF method side-by-side with the reliability of the comparable transaction method, given the use of those methods by financial analysts in the real world and the available evidence underlying each method.²⁷⁷

Determining the relative reliability of each methodology does not preclude the possibility that multiple methodologies may be reliably useful to decision making, particularly where multiple methodologies can serve to confirm the reliability of another. The investment tribunal in *CME v. Czech Republic*, for instance, used an adjusted DCF calculation, explained in considerable detail in the award, “as a confirmation of the Tribunal’s findings” based on an offer made by a potential buyer of the investment a six months before the treaty violation.²⁷⁸

Learning to Live with Uncertainty—Examining the relative degrees of reliability of competing methodologies may lead to an understanding that each has its own degree of uncertainty. But that examination does not lead inextricably to the conclusion that any amount of approximation renders a methodology speculative and unreliable. Many technical fields operate in full knowledge that conclusions are—necessarily—approximations and not unassailable facts.²⁷⁹ As many courts and tribunals in international disputes have recognized, uncertainty is an inextricable part of many technical fields, but the presence—and acknowledgment—of uncertainty in an expert’s conclusion does not mean that the methodology used to reach that conclusion is unreliable.²⁸⁰ At the same time, legal decision-makers are routinely asked to make deci-

276. See International Law Commission, *Draft Articles on Responsibility of States for Internationally Wrongful Acts* art. 36(2), 2001. Investment treaties generally provide little guidance on the issue. See Simmons, *supra* note 64, at 230 n.203.

277. If these tribunals had asked, they may have determined that their assumptions about the limits of DCF outside of the “going concern” context were misplaced. See Simmons, *supra* note 64, at 235 (“Financial analysts employ the method even for enterprises that arbitral tribunals would not deem ‘going concerns.’”).

278. See *CME Czech Republic B.V. v. Czech Republic*, Final Award, ¶ 604 (UNCITRAL Arb. Proc. 2003), <http://www.italaw.com/sites/default/files/case-documents/ita0180.pdf>.

279. See CATHERINE BUTTON, *THE POWER TO PROTECT: TRADE, HEALTH AND UNCERTAINTY IN THE WTO* 131 (2004) (“[T]he appropriate handling of uncertainties is part of the scientific process of risk assessment.”); Wirth, *supra* note 133, at 837 (“Because science is incomplete, the scientific data set underlying any regulation is necessarily incomplete. That, however, does not diminish the scientific nature of the inquiry.”).

280. See, e.g., *Himpurna Cal. Energy Ltd. v. PT (Persero) Persusahaan Listrik Negara (Indon.)*, Final Award, ¶ 376 (May 4, 1999), 25 Y.B. COMM. ARB. 13 (2000) (“There is no reason to apologise for the fact that [the DCF] approach involves approximations; they are inherent and inevitable. Nor can it be criticized as unrealistic or unbusinesslike; it is precisely how business executives must, and do, proceed when they evaluate a going concern. The fact that they use ranges and estimates does not imply abandonment of the discipline of economic analysis; nor, when adopted by arbitrators, does this method imply abandonment of the discipline of assessing the evidence before them.”); MARJORIE M. WHITEMAN, 3 *DAMAGES IN INTERNATIONAL LAW* 1694, 1699 (1943) (quoting *Delagoa Bay and East African Railway Company*, Award of 30 May 1900, regarding reliance on calculations of future income while

sions based on probabilities.²⁸¹ The focus should be on minimizing any potential error by examining the relative appropriateness of the methodology as compared to alternatives, the validity of the assumptions, and the reliability of the information underlying the expert's testimony.

Uncertainty can arise because the technical discipline is not sufficiently evolved to explain or describe the subject at issue, or uncertainty might arise simply because the necessary empirical evidence is unavailable.²⁸² And in every legal dispute, an expert's application of a methodology to a particular set of circumstances requires "interpretation of data," "conjectures on projected consequences," and individualized views on the "appropriate responses" to those data and consequences.²⁸³

What if an applicable legal or evidentiary rule demands a particular degree of certainty before an interest becomes legally protected or evidence is considered admissible?²⁸⁴ Sergey Ripinsky and Kevin Williams have posed this question in response to criticisms of the use of the DCF method investment arbitrations.²⁸⁵ They confront an apparent "contradiction" between "the two key principles in the law of damages: (a) the principle of prohibiting the award of speculative damages, and (b) the principle requiring the award of the 'fair market value' of the investment."²⁸⁶ As with other subjects of expert testimony, "[s]peculation and

acknowledging that "such a computation made in advance on the basis of purely theoretical data cannot hope to be absolutely accurate but only comparatively likely").

281. Alvarez, *supra* note 64, at 87–88 (citing investment disputes and international criminal cases in which "the tribunals were asked to make determinations involving probabilities and not verifiable certainties" and noting that "[d]eciding on the basis of uncertainty is what international and domestic courts do every day").

282. See THE CONSERVATION FOUNDATION, *RISK ASSESSMENT AND RISK CONTROL* 5 (1985)

Virtually all elements of risk assessment are clouded with uncertainty, basically of two kinds. First, the various scientific disciplines involved in assessing risk are not sufficiently developed either to explain the mechanisms by which particular causes produce particular effects or to provide good quantitative estimates of cause-and-effect relationships. Second, the data needed to analyze particular risks are usually not available.

See also Crawford-Brown, *supra* note 254, at 468 ("[U]ncertainty is an intrinsic and essential characteristic of science that must be passed on to the policy and legal arenas. Failure to do so, by insisting on a single estimate of risk . . . does not fully and truthfully describe the state of science at any moment.").

283. See Miles, *supra* note 33, at 157.

284. See, e.g., *Compania de Aguas del Aconquija, S.A. v. Argentine Republic*, ICSID Case No. ARB/97/3, Award, ¶ 8.3.4 (Aug. 20, 2007) (requiring future losses to be proved with "some level of certainty"); *Autopista Concesionada de Venezuela, C.A. v. Bolivarian Republic Of Venezuela*, ICSID Case No. ARB/00/5, Award, ¶ 351 (Sept. 23, 2003), (requiring future losses to prove "sufficient (degree of) certainty"); see also International Law Commission, *supra* note 276 (setting out a requirement that anticipated income streams to be proved with "sufficient certainty" to become a legally protected interest).

285. RIPINSKY & WILLIAMS, *supra* note 43, at 211.

286. See *id.*

uncertainty” is “inherent in any DCF analysis” but “can be dealt with through evaluating the constituent elements of the expert’s testimony, “by taking conservative estimates of cash flow projects and application of a higher discount rate.”²⁸⁷ Ripinsky and Williams ultimately conclude that a legal rule that demands “market” value necessarily contemplates a degree of uncertainty consistent with accepted practices in the relevant field, noting that, “[i]f one really wishes to estimate the *market* value of an investment, then one should use most common and accepted methods to reach that value, and the DCF method *is* an appropriate method.”²⁸⁸ Even if a decision maker determines that the applicable legal standard is fixed and cannot be reconciled with an expert’s methodology and its application to a dispute, it should still ask and answer the question of whether the relevant technical field would consider the expert evidence as reliable for the decision at hand.

Lastly, consensus among experts in the field or in decisional law on a methodology can evolve over time, and legal decision makers should not hesitate to take a fresh look at matters that were once considered settled, where the evidence suggesting a shifting expert consensus.²⁸⁹ In most technical fields, a dedication to empiricism means that what is considered generally accepted can and will change over time, and differences in views among specialists is likely to continue in varying degrees. When a particular hypothesis or assertion becomes generally accepted in a field notwithstanding those differences depends on the norms of the field, but in general occurs well before the hypothesis or assertion enters the unassailable “core” of knowledge in a field.²⁹⁰ One example is the choice between simple interest and compound interest in calculating monetary damages. Although it was perhaps once considered “settled” that compound interest was not allowable,²⁹¹ many more recent investment tribunals have considered it well-established that compound interest should be applied in calculating damages.²⁹²

287. *See id.*

288. *See id.*

289. *See Foster, supra* note 257, at 590 (“[T]he understanding that scientific knowledge is not static . . . is vital for any international tribunal dealing with a scientific case.”).

290. *See id.* at 10–11 (“What we know was an ‘invulnerable core of scientific knowledge ultimately consists of scientific claims that no scientist any longer challenges. . . . However, differences of view among scientists as to the validity of hypotheses and assertions can be expected to remain indefinitely in varying degrees.”).

291. *Compare* WHITEMAN, *supra* note 280, at 1997 (noting an apparent mid-century consensus); *see also* RIPINSKY & WILLIAMS, *supra* note 43, at 382. *But cf.* Natasha Affolder, *Awarding Compound Interest in International Arbitration*, 12 AM. REV. INT’L ARB. 45, 71–73 (2001) (“The authorities cited by Ms. Whiteman . . . fail to support the existence of a general principle of international law against the awarding of compound interest. At most, it can be said that the question of whether compound interest can be awarded is an unsettled question before international tribunals.”).

292. *See, e.g.,* Siag v. Arab Republic of Egypt, ICSID Case No. ARB/05/15, Award, ¶ 595 (May 11, 2009) (noting claimants’ submission that, “since 2000, no [fewer] than 15 out of 16 tribunals have awarded compound interest on damages in investment disputes”); *see also*

Relying on Available Data—The method that is most reliable in a particular case will depend in large part on the available data.²⁹³ In determining the fair market value of an investment, for example, experts have employed a variety of methodologies, including accounting for the “book value” of an investment, calculating the amount invested, assessing transactions comparable to the investment, calculating future cash flows of the investment according to the DCF method,²⁹⁴ or surveying a business’s sales history of expropriated inventory.²⁹⁵ The reliability of valuation methods that rely on prices actually paid for comparable property, for instance, depends significantly on the availability of recent and representative comparable sales.²⁹⁶ For that reason, tribunals have relied more readily on comparable sales when valuing real property, where representative comparable sales are more readily available, but have shied away from relying on such data when valuing entire businesses.²⁹⁷

5. Has the expert presented the evidence in a way that is useful in conveying information to a non-expert and clearly identifies key points of disagreement with other experts?

The predictability and legitimacy of legal decisions depends in large part on the transparency of the decision making process.²⁹⁸ But when it comes to expert evidence, decision makers do not bear the burden of transparency alone. The experts—who are involved in disputes to assist the decision maker in making a more well-informed and accurate decision—also have a role in ensuring that they present evidence in a transparent manner, consistent with the principles and practices of their field.²⁹⁹ A decision-maker should therefore demand that experts convey their

RIPINSKY & WILLIAMS, *supra* note 43, at 387 (describing compound interest as “com[ing] to be treated as the default solution”).

293. See MARBOE, *supra* note 64, at 205 (reasoning that the most reliable valuation method “depends . . . on the question of whether a tribunal finds that the main parameters of the method are appropriate and acceptable”).

294. See Simmons, *supra* note 64, at 223–24.

295. See MARBOE, *supra* note 64, at 195.

296. See *Kardassopoulos v. Republic of Georgia*, ICSID Case Nos. ARB/05/18, ARB/07/15, Award, ¶ 598 (2010) (“It is not common in investment treaty arbitrations that a Tribunal has available to it three arm’s-length, contemporaneous transactions (or potential transactions) to assist in valuing an investment, much less three that converge in a narrow range of value.”); see also *Phelps Dodge v. Iran*, 10 Iran-U.S. Cl. Trib. Rep. 121, ¶¶ 29–30 (1986) (rejecting valuation by multiples because the consultant’s study relied on U.S. enterprises that were not sufficiently comparable to a start-up business in Iran).

297. See MARBOE, *supra* note 64, at 201–02 (noting that “the Iran-US Claims Tribunal frequently employed special experts in order to provide information about prices actually paid in the real property market” but that “the comparable sales approach has not been applied with respect to entire investment projects or undertakings in international investment disputes”).

298. See Walker, *supra* note 61, at 165.

299. See *id.* at 165–66 (noting that with respect to transparency “law and science share the same ideal”).

opinions in a way that is readily accessible and understandable to non-experts such as the decision-makers and future readers of the decision maker's conclusions.³⁰⁰

As an initial matter, a decision maker should ensure that each expert has clearly disclosed all key assumptions and facts underlying the evidence. This is the first step in understanding how an expert has applied the chosen methodologies to the specifics of the case at hand. It also helps to counteract an inherent asymmetry between experts, who know and understand a methodology's inputs, and non-experts, who are unlikely to know and understand those inputs without disclosure and explanation.³⁰¹ Many evidentiary guidelines and rules now provide that experts should make such disclosures. The IBA rules, for example, provide that an expert should disclose a statement of the facts on which he or she is basing his or her opinions and a description of the method, evidence, and information used in arriving at the expert's conclusion.³⁰² Mark Kantor has proposed that investment arbitrators impose on party-appointed experts a duty to provide "full information" to the adjudicators, whether that information supports or contradicts the expert's professional analyses and conclusions, and a duty to use assess, to the extent an expert has the background to do so, the reasonableness of assumptions provided by counsel or a party.³⁰³ And Joost Pauwelyn has argued that parties relying on empirical economic data should disclose all assumptions and that adjudicators should transparently evaluate those assumptions in their decisions.³⁰⁴

In this regard, one indicator that an expert has correctly applied a methodology is disclosure of a transparent and flexible roadmap of underlying assumptions so that a tribunal may choose which of an expert's assumptions to accept and which to reject.³⁰⁵ Take, for example, the expert

300. See Foster, *supra* note 3, at xiv ("The science must be put forward in a form that is readily digestible by a court or tribunal composed of individuals whose qualifications and experience lie in the field of law rather than science.").

301. Cf. Lorna Schrefler, *Reflections on the Different Roles of Expertise in Regulatory Policy Making*, in *THE ROLE OF "EXPERTS" IN INTERNATIONAL AND EUROPEAN DECISION-MAKING PROCESSES: ADVISORS, DECISION MAKERS OR IRRELEVANT ACTORS?* 63, 69–70 (M. Ambrus et al. eds., 2014) ("Another example could be when data or assumptions in an economic analysis are carefully selected to present a certain picture of reality, often exploiting the information asymmetries between agents (in this case, the experts within the [regulatory or administrative] agency) and principals.").

302. IBA RULES 5.3(d)-(e); see also Wisner et al., *supra* note 94, at 243–44.

303. Mark Kantor, *A Code of Conduct for Party-Appointed Experts in International Arbitration – Can One Be Found?*, 26 *ARB. INT'L* 323, 375 (2010).

304. Pauwelyn, *supra* note 77, at 187 ("Assumptions [underlying economic empirical methodologies] must be disclosed and discussed. . .").

305. See *Starrett Housing Corp. v. Iran*, Final Award, 16 Iran-U.S. Cl. Trib. Rep. 112, ¶ 39 (1987) ("[T]he Expert referred many of his assumptions and decisions with regard to his valuation to the Tribunal . . . [in order to] ma[k]e it easier for the Tribunal to understand and possibly to adjust the final result according to the Tribunal's own judgement."); *id.* ¶ 269 ("[W]here [the expert] drew inferences or made subjective judgments, he pointed them out and explained his reasons."); see also *Corfu Channel*, (U.K. v. Alb.), Order, 1948 I.C.J. Rep. 124, 126–27 (Dec. 17) (instructing a panel of three experts to "give the reasons for these

report in the *Manica Plateau Arbitration*. The report first laid out—in plain language—the technical terms at the core of the parties’ dispute, and how they were understood in geographical science.³⁰⁶ The expert then set out to apply these “geographical principles” to the facts of the dispute and the arguments of the parties, noting where the principles pointed one way or the other.³⁰⁷ Only then did the expert turn to answering the specific questions put to him by the arbitrator, laying out each incremental conclusion by reference to the previously defined principles and analysis.³⁰⁸ By setting out the relevant principles, assumptions, and facts, and connecting them to his conclusions, the expert permitted the arbitrator to assess each distinct line of analysis and modify as necessary.³⁰⁹

Outside the dispute context, experts commonly account for uncertainty in their conclusions through statistical methods such as sensitivity analyses, which describe how sensitive an expert’s conclusions are to changes in underlying data, methodology, or assumptions. In disputes, they also can assist decision makers in evaluating a range of inputs and outcomes.³¹⁰ This is especially true where small changes in assumptions and other inputs can significantly alter an expert’s conclusions.³¹¹ A reliable expert witness will disclose those assumptions and inputs, explain how they drive the analysis, and identify material differences between that analysis and the analyses of the other party’s experts. In *Guarachi America, Inc. & Rurelec PLC v. Bolivia*, for example, the claimants’ valuation expert provided an assessment that “95% of the difference” between the parties’ valuations was due to three issues: discount rate, revenue pro-

findings in order to make their true significance apparent to the Court. . . . [and] mention any doubts or differences of opinion amongst them”).

306. *Manica Plateau Arbitration* (Gr. Brit. v. Portugal), 28 R.I.A.A. 283, 302 (1897).

307. *Id.* at 302–03.

308. *Id.* at 303–07.

309. *Id.* at 308–09. The arbitrator reasoned that, although the expert’s conclusions were “technically accurate,” the result was so “irregular” and full of “numerous inflections,” that it could “easily give rise . . . to doubts and differences of opinion which should be carefully avoided.” *Id.* at 308. As a result, the arbitrator asked the expert to modify one segment of the proposed demarcation to “substitut[e] some nearly straight and better-defined lines for the natural inflections . . . so that the extent of ground which each party gets by the substitution of straight lines for the rigorous demarcation of the edge remains almost equivalent.” *Id.*

310. *See* *Starrett Housing*, 16 Iran-U.S. Cl. Rep. ¶¶ 39–40 (1987) (noting that tribunal-appointed valuation expert “provided a sensitivity analysis for four essential areas to show what effect their alteration would have on the valuation”); Pauwelyn, *supra* note 77, at 187 (suggesting that adjudicators require “thorough robustness analysis” of economic data used in international disputes).

311. *See* *Land Reclamation by Singapore in and Around the Straits of Johor* (Malay. v. Sing.), Verbatim Records, 37–38 (Sept. 25, 2003 a.m.), https://www.itlos.org/fileadmin/itlos/documents/cases/case_no_12/PV.03.01.25.09.03.a.m.E.pdf (eliciting expert testimony that the sediment transport rates at the core of the land reclamation dispute could change significantly in response to only slight changes in velocity).

jections, and capital expenditures.³¹² The tribunal noted on several occasions that each expert had disclosed critical assumptions, enabling the other party's experts—and the tribunal itself—to confirm or challenge them.³¹³ By contrast, an expert that requires a decision maker to accept all underlying assumptions in order to rely on the testimony is unlikely to provide valuable assistance and rightfully risks flat rejection.³¹⁴

6. *Was the methodology applied within practical and theoretical boundaries?*

It is of course not enough to determine that a method is reliable in the abstract or that a witness is a legitimate expert in the methodology. A tribunal must also ensure that the expert has reliably applied the methodology to the particular circumstances of the dispute at hand.³¹⁵ This generally requires identifying the key premises and limitations of an expert's methodology and determining whether the assumptions and inputs that have gone into the expert's analysis are sound.

With respect to the key premises and limitations of an expert's methodology, a decision maker should confirm that the expert has applied the methodology consistently, mindful of the methodology's key premises and limitations identified earlier in the reliability analysis. Reliable expert evidence will not cherry-pick elements of a methodology that favor the expert's conclusions but will instead faithfully apply the methodology within its theoretical and practical limitations. If an expert has applied a peer-reviewed methodology, citing certain publications or other sources, what methodological limits do those sources recognize? What limitations do other peer-reviewed sources recognize?

With respect to the assumptions and inputs underlying expert opinion, a decision maker should ask whether they would be considered sufficiently reliable in the relevant technical field³¹⁶ and whether the expert has attempted to verify all information and assumptions provided by the parties

312. *Guarachi America, Inc. v. Plurinational State of Bolivia*, Award, PCA Case No. 2011-17, ¶ 454 (Jan. 31, 2013).

313. *See, e.g., id.* ¶¶ 505–510 (noting differences in the experts' assumptions regarding future capacity of the investment, including assumptions related to demand and competition in the market).

314. *See Wisner et al., supra* note 94, at 247–48 (citing an example of expert testimony that permitted the decision maker to accept or reject each key assumption used in the expert's valuation).

315. *See RIPINSKY & WILLIAMS, supra* note 43, at 194 (“[E]very valuation requires careful analysis specific to the circumstances of the case.”).

316. *See MARBOE, supra* note 64, at 219 (emphasizing that the underlying “forecast of future cash flows” in a DCF analysis must be “checked for plausibility in order to determine its reasonableness and lack of contradiction”); Pauwelyn, *supra* note 77, at 187 (“Assumptions must . . . be consistent with the particular product, industry or market under examination.”); *cf.* Orellana, *supra* note 180, at 54 (noting that international adjudicators can evaluate the reliability of scientific conclusions without deciding questions of “scientific validities” by “resort[ing] to established techniques in fact-finding” and examining underlying evidence).

or the tribunal or court.³¹⁷ If they are not, testimony that relies on such data is not entitled to much, if any, weight.³¹⁸

The tribunal's analysis in *Guarachi America, Inc. & Rurelec PLC v. Bolivia* provides a particularly useful illustration of this basic approach.³¹⁹ The tribunal in that dispute expressly divided its analysis of the parties' expert damages evidence into three primary areas of disagreement, identified the assumptions and inputs driving that disagreement, and compared each assumption and input against the available documentary evidence.³²⁰ The result was a thoroughly reasoned valuation that sought to adopt the most reliable assumptions and inputs from both parties' experts and other evidence in the record.³²¹ The tribunal also looked to the primary sources on which the experts relied to ensure that the methodology had been applied within its generally accepted boundaries. For instance, when Bolivia's expert sought to apply a methodology that would have increased the country risk associated with the investment, thereby increasing the discount rate and lowering the valuation, the tribunal went back to the primary sources on which the expert had relied and discovered that those sources would apply the methodology only in short-term scenarios (e.g., five to 10 years), not in the longer-term time period (28 years) at issue in *Guarachi America*.³²²

When using the DCF method the assumptions and inputs critical to the reliability of the valuation methodology can include past performance, internal business plans and contract rights, and the regulatory framework

317. See *Starrett Housing Corp. v. Iran*, Award, 16 Iran-U.S. Cl. Rep. 112, ¶ 268 (1987) (reasoning that it gave "substantial weight" to an expert report in part because of the "thoroughness of the process by which the Expert sought to verify all information presented to him by the Parties").

318. See, e.g., *Arbitration Under Article 181 of the Treaty of Neuilly* ("Certain Forests in Rhodopia"), Decision on the Principal Question (Mar. 29, 1933), 28 AM. J. INT'L L. 760, 805 (1934). In *Certain Forests in Rhodopia*, Greece alleged that Bulgaria had dispossessed Greek nationals of their rights in forest land during World War I. *Id.* at 760. The arbitrator declined Bulgaria's request for an expert assessment of the value of timber, reasoning that the only evidence on which an expert could have relied – an in-person survey of the timber – would have resulted in "speculative" testimony, in part because an expert could not have taken into account any logging activities that had occurred in the 15 years since the alleged dispossession. See *id.* at 805

The Arbitrator does not believe that it would be practical at the present time to proceed with such a survey, considering the length of time that has elapsed since the date of the seizure of the forests by Bulgaria, and considering the fact that considerable cutting may have been done in that interval for the benefit of other persons.

319. See *Guarachi America, Inc. v. Plurinational State of Bolivia*, Award, PCA Case No. 2011-17 (Jan. 31, 2013).

320. See *id.* ¶¶ 452–617.

321. See, e.g., *id.* ¶¶ 505–510.

322. See *id.* ¶¶ 571–581 (noting that the sources cited by Bolivia's expert made clear that the expert's "multiplier" approach, which would have increased the country risk associated with the investment at issue based on broad macroeconomic risks in emerging markets such as Bolivia, was appropriate only in short-term valuations).

and larger macroeconomic climate.³²³ The tribunal in *CMS v. Argentina*, for example, analyzed the key assumptions of the claimant's valuation expert, making a number of modifications to the valuation model.³²⁴ Of particular importance, the tribunal reasoned that in determining future cash flows from the investor's natural gas transmission business the expert had erred in assuming that Argentina's economic crisis would not have reduced demand for gas.³²⁵ Similarly, in valuing the claimant's investment in an oil production project, the tribunal in *Venezuela Holdings v. Venezuela* relied on the claimant's oil price forecast because it included all relevant information available at the time of the valuation date, including a decision by the Organization of the Petroleum Exporting Countries not to increase production, which would have resulted in an increase in supply and likely decrease in price.³²⁶

Although a reliability analysis of expert evidence must to some degree include an examination of the constituent parts of an expert's opinion, a legal decision-maker should in the end consider an expert's presentation as a whole and the totality of the evidence, and determine whether the "weight of the evidence" supports a conclusion that the expert evidence is reliable.³²⁷ Under this approach, a decision maker should not discard expert evidence merely because some aspects are potentially flawed but should examine the evidence as a whole, taking into account any flaws in that broader context. In *CMS v. Argentina*, for example, the tribunal noted that the "general approach" under the DCF method of the claimant's valuation expert "remain[ed] . . . valid," even though the tribunal the tribunal expressed a number of "reservations" with the expert's application of the approach and made "a number of changes" to the expert's assumptions.³²⁸

CONCLUSION

Learned Hand wrote in 1901 that "[n]o one will deny that the law should in some way effectively use expert knowledge wherever it will aid in settling disputes," noting that "[t]he only question is as to how it can do

323. See MARBOE, *supra* note 64, at 219-28; see also CME Czech Republic B.V. v. Czech Republic, Final Award, ¶¶ 554-566, (UNCITRAL Arb. Proc. 2003), <http://www.italaw.com/sites/default/files/case-documents/ita0180.pdf> (examining the factors influencing an expert's valuation and making adjustments where "certain deductions" were "unsustainable").

324. See CMS Gas Transmission Company v. Argentine Republic, ICSID Case No. ARB/01/8, Award, ¶¶ 435-463 (May 12, 2005).

325. See *id.* ¶ 444.

326. See Venezuela Holdings B.V. v. Bolivarian Republic of Venezuela, ICSID Case No. ARB/07/27, Award, ¶ 327 (2015).

327. See MCGARITY & WAGNER, *supra* note 225, at 154-56; see also BROWER & BRUESCHKE, *supra* note 90, at 201 (noting that arbitral tribunals "must determine what weight, if any, based upon the totality of the evidence, to give to the expert's conclusions").

328. See CMS Gas Transmission Company, ICSID Case No. ARB/01/8, Award, ¶ 434.

so best.”³²⁹ More than a century later, that question still remains, and international law continues to evolve to more effectively use expert evidence. The question also grows more complex by the day, as human activity, the disputes that arise from that activity, and the international agreements that set out rules for addressing those disputes, also grow in scope and their own complexity.

This Article attempts to complement existing efforts to better analyze expert evidence, with a focus on the reliability of an expert’s methodology and its application to a particular set of facts. A more robust, comprehensive, and transparent framework built on best practices developed within international law may assist in reaching more analytically sound decisions and enhance legitimacy.

329. Learned Hand, *Historical and Practical Considerations Regarding Expert Testimony*, 15 HARVARD L. REV. 40, 40 (1901).