

# CONSTRUCTING MORE RELIABLE LAW AND POLICY: THE POTENTIAL BENEFITS OF THE UNDERUSED DELPHI METHOD

Juan Bataller-Grau, Elies Seguí-Mas, Javier Vercher-Moll and Jeffrey W. Stempel\*

## I. INTRODUCTION: THE UNDULY AD HOC NATURE OF LEGAL DECISION-MAKING AND THE NEED FOR GREATER INFUSION OF EXPERTISE

The legal system in both the United States and Europe, in both common law and civil law countries, is often criticized as needing to be more “scientific” regarding factual inquiry, empiricism, technology and scientific principles.<sup>1</sup> Although these concerns-cum-criticisms may be overstated,<sup>2</sup> they contain a

---

\* Respectively, Professor of Business Law, Universidad de Valencia, Associate Professor of Economics, Universitat Politecnica de Valencia, Lecturer, Universidad de Valencia, and Doris S. & Theodore B. Lee Professor of Law, William S. Boyd School of Law, University of Nevada–Las Vegas. The authors thank M<sup>a</sup> Jose Añón, Vicente Cuñat, Sara Gordon, Javier de Lucas, Dan Hamilton, Helmut Heiss, Ted Lee, Ann McGinley, David McClure, Martin Schauer, Manfred Wandt and our colleagues in the Principles of Reinsurance Contract Law Project Group. © 2019 Juan Bataller- Grau, Elies Seguí- Mas, Javier Vercher-Moll & Jeffrey W. Stempel 2019.

\*\* To preserve author voice and originality, *UMKC Law Review* has maintained the author’s original spelling and word choices, including the titles of Spanish-language sources, where there will be no resulting confusion.

<sup>1</sup> See generally LAW AND SCIENCE: VOLS. I & II (Susan Silbey ed., 2008); THE USE/NON USE/MISUSE OF APPLIED SOCIAL RESEARCH IN THE COURTS (Michael Saks & Charles Baron eds., 1980); Joshua B. Fischman, *Measuring Inconsistency, Indeterminacy, and Error in Adjudication*, 16 AM. L. & ECON. REV. 40 (2013); John McCoid, II, *Inconsistent Judgments*, 48 WASH. & LEE L. REV. 487 (1991); Fowler Harper, *Scientific Method in the Application of Law*, 1 DAKOTA L. REV. 110 (1927); Roscoe Pound, *Need for Sociological Jurisprudence*, 19 GREEN BAG 607 (1907). See also David Cole, *Against Literalism*, 40 STAN. L. REV. 545 (1988) (reviewing JAMES B. WHITE, HERACLES BOW: ESSAYS ON THE RHETORIC AND POETICS OF LAW (1985)) (Book Review) (“In legal theory, a little fact goes a long way. From the recognition that law is transmitted through words springs the school of ‘law and literature.’ From the fact that legal disputes involve the allocation of costs comes ‘law and economics.’”).

<sup>2</sup> For example, regarding what might be termed individual or situation-specific fact such as the events leading to a vehicle collision, the origin of a fire, or building collapse, traditional trials appear to be very good at determining the cause of the problem and the extent of damage. The trial process is less good at determining matter outside the immediate context of a particularized dispute. Judges and jurors have their own impressions of the world at large but these impressions are generally not tested or developed by the trial process. The legislative process is supposed to make policy, but often does so on the basis of political imperatives or expediency rather than sustained examination of a problem, with much of the information received from lobbyists or interest groups. See, e.g., Editorial, *Marsy’s Law Would Create More Victims in Florida’s Courts*, ORLANDO SUN-SENTINEL (Mar. 19, 2018), <http://www.sun-sentinal.com/opinion/editorials/fl-op-editorial-marsys-law-proposal-96-story.html> (noting that proposed victim’s rights law, popular in the abstract, would have unintended consequences hindering law enforcement, such as “allow[ing] victims to bypass the prosecution and petition judges for hasty trials, as well as set arbitrary limits on the time to complete a state-level appeal” as well as defining victimhood so broadly to vest rights in, for example, all students in a school victimized by mass shooting, regardless of injury or proximity to danger). In the United States, administrative agencies were supposed to incorporate expertise into policy-making and application but have fallen short of expectations due to agency capture and political intrusion. See Daniel

substantial kernel of truth: courts can render disparate outcomes in seemingly similar disputes involving similar products, processes, procedures, conduct, and injury.<sup>3</sup> Legislation may be contradictory, as may executive orders or administrative agency decisions. Some of this is, of course, the natural result of the different contexts and decision makers as well as different trial settings and (in the United States) use of lay juries. But in adjudication, judges dealing with the same activities, laws, public policy, or contract language often reach disparate results that suggest undue inconsistency and unpredictability in the law.

Even when holdings are consistent, the judiciary is occasionally embarrassed when courts make broad factual or scientific pronouncements that are just plain false or based on outdated, mis-cited experiments, or publications that have subsequently been debunked.<sup>4</sup> Sometimes judges display a complete misunderstanding of actual facts concerning business, science, history, sociology,

---

Rodrigues & Barry Weingast, *The "Reformation of Administrative Law" Revisited*, 31 J. L. ECON. & ORG. 782 (2015); Jim Rossi, *Participation Run Amok: The Costs of Mass Participation for Deliberative Agency Decisionmaking*, 92 NW. U. L. REV. 173 (1997); Richard Stewart, *The Reformation of American Administrative Law*, 88 HARV. L. REV. 1669 (1975) (all noting degree to which agencies become identified with support for industries they were supposed to regulate with dispassion and degree to which elected official spurred by interest groups may intervene with agency's objective decision-making, problems that may be exacerbated by agency turnover but differing as to role of courts in reviewing agency action and degree to which decision-making in the public interest has been undermined by the activity of regulatory constituents). *See also* Richard Stewart, *Administrative Law in the Twenty-First Century*, 78 N.Y.U. L. REV. 437 (2003) (taking more optimistic view of administrative regulation in its primary forms and describing rulemaking in terms compatible with Delphi Method injection of expertise into agency analysis).

<sup>3</sup> *See generally* Kyle Tsui & Buchanan PTAB Report, *Federal Circuit Holds that the PTAB and District Courts May Reach Different Conclusions*, BUCHANAN PTAB REPORT, INGERSOLL & ROONEY PC (Apr. 11, 2017), <http://buchanan-ip.com/PTAB/federal-circuit-holds-ptab-district-courts-may-reach-different-conclusions/> (finding that district courts and the Patent Trial and Appeal Board may reach contrary conclusions on the validity of challenged patents even when presented with essentially the same evidence).

<sup>4</sup> *See* Allison Orr Larsen, *The Trouble With Amicus Facts*, 100 VA. L. REV. 1757 (2014) (noting that many amicus briefs submitted to the U.S. Supreme Court contained false or unsubstantiated assertions of empirical fact and that these unsupported or even blatantly incorrect assertions sometimes are embraced by the Court without sufficient vetting); John Pfaff, Opinion, *The Supreme Court Justices Need Fact-Checkers*, N.Y. TIMES (Oct. 18, 2017), available at <http://nyti.ms/2kXEVDH>. *See, e.g.*, *McKune v. Lile*, 536 U.S. 24 (2002), discussed in the ensuing footnote, *infra* note 5. *See also* Allison Orr Larsen, *Confronting Supreme Court Fact Finding*, 98 VA. L. REV. 1255 (2012) (finding risk not only in uncritical acceptance of erroneous facts asserted to the court by interested parties but also by the Court's own internal research not subject to transparency and verification).

economics, or statistics.<sup>5</sup> The same can be true of legislators<sup>6</sup> or administrative agencies.<sup>7</sup>

---

<sup>5</sup> See, e.g., *McKune v. Lile*, 546 U.S. 24 (2002). *McKune v. Lile* provides a particularly embarrassing illustration for the Court. Picking up on distortions contained in an amicus brief favored by the Court, the majority opinion stated that “[t]he rate of recidivism of treated sex offenders is fairly consistently estimated to be around 15%, whereas the rate of recidivism of untreated offenders has been estimated to be as high as 80%.” *Id.* at 33 (citing Robert Freeman-Longo and Robert Wall, *Changing a Lifetime of Sexual Crime*, *PSYCHOLOGY TODAY* 58-64 (1986)). The actual rate of recidivism among sex crimes offenders is a mere 1.5%. See U.S. DEPT. OF JUST., BUREAU OF JUST. STAT., *RECIDIVISM OF PRISONERS RELEASED IN 30 STATES IN 2005, BY MOST SERIOUS COMMITMENT OFFENSE AND TYPES OF POST-RELEASE ARREST CHARGES* tbl. 2 (2014), <http://www.bjs.gov/content/pub/pdf/rprts05p0510.pdf>. Professor Freeman-Longo also has stated that the *Psychology Today* article misstated his research. See Ira Mark Ellman & Tara Ellman, *Frightening and High: The Supreme Court’s Crucial Mistake About Sex Crime Statistics*, 30 *CONS. COMM.* 495 (2015) (noting that the erroneous 80% statistic had been cited by more than 90 lower courts); Radley Balko, Opinion, *How A Dubious Statistic Convinced U.S. Courts to Approve of Indefinite Detention*, *WASH. POST* (Aug. 20, 2015), <http://wapo.st/1NxOxB1>. See EXEC. OFFICE OF THE PRESIDENT, PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE & TECH. (“PCAST”), *REPORT TO THE PRESIDENT, FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS* (2016) (debunking or casting doubt upon reliability of several frequently admitted types of forensic evidence, including fingerprints, basic evidence, and bite marks) [hereinafter PCAST REPORT]. Joelle Anne Moreno, *Einstein on the Bench? Exposing What Judges Do Not Know About Science and Using Child Abuse Cases to Improve How Courts Evaluate Scientific Evidence*, 64 *OHIO ST. L.J.* 531 (2003); J. Alexander Tanford, *The Limits of a Scientific Jurisprudence: The Supreme Court and Psychology*, 66 *IND. L.J.* 137 (1990); Michael J. Saks, *Ignorance of Science is No Excuse*, 10 *TRIAL* 18 (1970) (searing criticism of U.S. Supreme Court decisions expressing no concern about jury size reduction from twelve to six jurors but finding constitutional minimum of five jurors on the basis of research suggesting reduction to six was equally problematic). See also *STATISTICAL SCIENCE IN THE COURTROOM* (Joseph Gastwirth ed., 2000); Alisa Smith & Michael J. Saks, *The Case for Overturning Williams v. Florida and the Six-Person Jury: History, Law and Empirical Evidence*, 60 *FLA. L. REV.* 441 (2008); Michael J. Saks, *Protecting Factfinders from Being Overly Misled, While Still Admitting Weakly Supported Forensic Science into Evidence*, 43 *TULSA L. REV.* 609 (2008).

<sup>6</sup> For example, legislators in Spain have repeatedly reacted to dramatic losses by requiring compulsory insurance to cover such losses without any analysis of whether insurers are willing or even able to provide such insurance products at a profit. As a consequence, there are now more than fifty types of insurance that are compulsory in Spain—at least in theory and according to the letter of the law. In practice, comparatively few companies offer many of these coverages and comparatively few persons or entities purchase these coverages. See also William D. Araiza, *Deference to Congressional Fact-Finding in Rights-Enforcing and Rights-Limiting Legislation*, 88 *NYU L. REV.* 878 (2013) (crediting Congress with great capacity for empirical fact-finding but suggesting that judiciary should be less deferential to legislative determination of such empirical facts than courts are to legislative determination of value-based facts, sometimes referred to as “legislative facts”); Kenneth Prewitt, *The Congressional War on the Social Sciences*, *PACIFIC STANDARD* (May 24, 2013), <http://www.psmag.com/news/the-congressional-war-on-social-science-58407>. The typology of legislative (or value-based) facts and “adjudicative” or empirical fact is set forth in the classic article by Kenneth Culp Davis, *An Approach to Problems of Evidence in the Administrative Process*, 55 *HARV. L. REV.* 364, 402 (1942). In our view, the typology and different levels of judicial deference makes basic sense. Courts should be generally more deferential to legislatures when their “fact-finding” smacks more of expressing preferences (e.g., the legislature “finds that consumers will benefit from greater competition and thus benefit from a larger number of vendors in the industry”) than stating empirical, real-world facts (e.g., a congressional report claiming that a certain number of

Law has long aspired to achieve status as a science. A central theme of much legal philosophy has been the quest for legal doctrine to become more like scientific axioms or findings produced through a scientific inquiry.<sup>8</sup> Considerable debate has surrounded the issue. Part of the legal profession sees the question of law's science status as doomed to failure and regards law as a distinct type of discipline.<sup>9</sup> Others in the legal profession are attracted to the aspiration but express

---

undocumented workers collect a specific amount of public assistance in a given year). If a legislative body, as demonstrated by examination of reputable sources, is wrong about the latter assertion of empirical facts, no deference is owed. Gross errors of empirical fact-finding (or gross assertions of empirical fact that are incorrect) may even be treated as the metaphorical "thirteenth chime of the clock" that supports skepticism about the entire legislation surrounding errors in fact-finding, particularly if an error has its source in the representations of interested lobbyists. For purposes of this article and as explained later in the article, we note that use of the Delphi Method can greatly reduce the risk of legislative errors of empirical fact by self-consciously involving experts in the legislative process. Value-based fact finding may also be improved as well because the Delphi process is more likely than ad hoc methods to ensure that different values and ideologies are sufficiently represented in the process.

<sup>7</sup> For example, Spanish insurance regulators in June 1996 took the position that liability insurance should be prohibited for government entities on the theory that the existence of such insurance would create undue moral hazard and make government employees less careful in discharge of their duties. Faced with the reality of possible liability exposure, even the relatively limited exposure of the sovereign, governments in practice continued to purchase liability insurance from private insurers. See also Theodore J. St. Antoine, *The NLRB, the Courts, the Administrative Procedure Act, and Chevron: Now and Then*, 64 EMORY L.J. 1529 (2015) (noting degree to which despite agency claims of objective expertise, empirical study of the National Labor Relations Board indicates that "the political backgrounds of NLRB members substantially affect its decisions" including fact finding and the inferences drawn from facts, which in turn raises "fundamental questions about federal neutrality in union-management relations")(italics removed). In our view, such findings provide support for a more structured expert examination of the factual environment of labor disputes by a cross-section of experts such as is provided by the Delphi Method. In some cases, an error made by the agency may be caught, disclosed, and correct. See Alexander Merritt, *Confession of Error by Administrative Agencies*, 67 WASH. & LEE L. REV. 1197 (2010).

<sup>8</sup> See William Schofeld, *Christopher Columbus Langdell*, 55 U. PA. L. REV. 273, 281 (1907) (noting that famed Harvard Law Dean credited with establishing the Case Method and the Socratic Method as law teaching vehicles "asserted and believed that law is a science" and "that there is a scientific method of teaching and studying the law."). See also *id.* at 281-82 ("His whole endeavor was to introduce and establish that scientific method.").

<sup>9</sup> See generally ROBERT B. STEVENS, *LAW SCHOOL: LEGAL EDUCATION IN AMERICA FROM THE 1850S TO THE 1980S* 39 (1983) (discussing efforts of prominent Harvard Law Dean Christopher Langdell to classify law as a science and subsequent argument that law is more art than science or is a distinct discipline not readily classified as science). Although the view that law is a unique discipline still continues to exert a considerable hold over the legal profession, there is substantial ground for concluding that law has lost some of its status as a self-contained discipline or even a particularly exclusive profession. See, e.g., BRIAN Z. TAMANAHA, *FAILING LAW SCHOOLS* (2012); PAUL F. CAMPOS, *DON'T GO TO LAW SCHOOL (UNLESS)* (2012); RICHARD SUSSKIND, *THE END OF LAWYERS? RETHINKING THE NATURE OF LEGAL SERVICES* (2008); William D. Henderson, *Three Generations of U.S. Lawyers: Generalists, Specialists, Project Managers*, 70 MD. L. REV. 1 (2011); Herbert M. Kritzer, *The Commodification of Insurance Defense Practice*, 509 VAND. L. REV. 2053 (2006); Russell G. Pearce, *The Professionalism Paradigm Shift: Why Discarding Professional Ideology Will Improve the Conduct and Reputation of the Bar*, 70 NYU L. REV. 1229 (1995); Richard A. Posner, *The Decline of Law as an Autonomous Discipline 1962-1987*, 100 HARV. L. REV. 761 (1987). Adding

doubt regarding whether the methods that the legal doctrine has traditionally employed can achieve the greater apparent rigor of the physical sciences, or even the social sciences.<sup>10</sup>

One tool for at least making strides toward a more scientific brand of law is the Delphi Method, a regime for capturing expert analysis developed by the Rand Corporation during the mid-twentieth century.<sup>11</sup> Judicious injection of the Delphi Method principles and practices can make both legal policymaking and application of the law more consistent, systematic, reflective, consistent and wise—more “scientific,” if you will. We are not proposing a change in the law’s paradigm but rather proposing that the law-making and the legal system make self-conscious use of the Delphi Method in apt situations. A too-often overlooked methodology can play a relevant role in law making, legal research, and adjudication.

## II. AN OVERVIEW OF THE DELPHI METHOD AND ITS SEEMING FADE FROM ACADEMIC LITERATURE

The Delphi Method has been summarized as a structured methodology for using the systematic interaction of a panel of experts as a means of making determinations and forecasts.<sup>12</sup> The Method’s origin as a means of forecasting outcomes is derived from the Greek Oracle of Delphi, who foretold the future.<sup>13</sup> More prosaically, the Method grew out of World War II and the Cold War when a report was ordered in 1944 for the U.S. Army Air Corps addressing the potential future military application of new technologies.<sup>14</sup> After experimenting with several

---

to challenges presented by changes in what might be termed the traditional face-to-face market for legal services has been the rise of technological substitutes for a human lawyer such as Quicken or LegalZoom.

<sup>10</sup> See, e.g., Thomas S. Ulen, *A Nobel Prize in Legal Science: Theory, Empirical Work and the Scientific Method in the Study of Law*, 2002 U. ILL. L. REV. 875, 878 n.8 (law properly classified is a social science); see also Rolien M.C. Roos, *Is Law a Science?* 17 P.E.L.J./P.E.R. 4 (2014); George Goble, *Law as a Science*, 9 IND. L.J. 294 (1934) (taking position that law is a science that can be pursued through scientific methodology using judicial decisions as data points).

<sup>11</sup> See generally *Delphi Method*, RAND CORP., <http://www.rand.org/topics/delphi-method.html> (last visited Feb. 1, 2019); see also OLAF HELMER, RAND PAPER NO. P-3558., ANALYSIS OF THE FUTURE: THE DELPHI METHOD (1967). For an extensive review of the Delphi Method, see THE DELPHI METHOD: TECHNIQUES AND APPLICATIONS (Harold A. Linstone & Murray Turoff eds., 2002).

<sup>12</sup> See Kesten C. Green et al., *Methods to Elicit Forecasts from Groups: Delphi and Prediction Markets Compared*, 8 FORESIGHT: INT’L J. APPLIED FORECASTING 1 (2007) (concluding that Delphi Method is effective but under-utilized); J. Scott Armstrong, *Comparing Face-to-Face Meetings, Nominal Groups, Delphi and Prediction Markets on an Estimation Task*, 27 INT’L J. FORECASTING 183 (2011) (finding use of Delphi Method has relatively greater predictive value than other methods).

<sup>13</sup> The nomenclature is a bit unfortunate both because it implies a false sense of certainty about Delphi Method analysis and was disliked by some of its originators as “smacking a little of the occult.” See *Delphi Method*, INVESTOPEDIA, <http://www.investopedia.com/terms/d/delphi-method.asp> (last updated Aug. 7, 2018); see also Green et al., *supra* note 12.

<sup>14</sup> See NICHOLAS RESCHER, PREDICTING THE FUTURE: AN INTRODUCTION TO THE THEORY OF

approaches, military contractor Rand Corporation developed Delphi during the 1950s.<sup>15</sup> After it became available in the public domain during the 1960s, Delphi acquired a following, particularly in organizational and business planning.<sup>16</sup>

But Delphi never really caught on in law, and by the late 1970s it had generally faded as an academic topic altogether,<sup>17</sup> and has been less frequently discussed (but not necessarily less used “below the surface”) by businesses and governments since that time.<sup>18</sup> Outside the legal field, Delphi may have faded from obvious view, but it remains in use and discussion. For example, Delphi advocates have argued the methodology is particularly well-suited to obtaining expert input

---

FORECASTING (1998).

<sup>15</sup> Olaf Helmer, Norman Dalkey and Nicholas Rescher are generally credited as the creators of Delphi. *See, e.g.*, Norman Dalkey & Olaf Helmer, *An Experimental Application of the Delphi Method to the Use of Experts*, 9 MGMT. SCI. 458 (1963); *see also* BERNICE B. BROWN, RAND PAPER NO. P-3925, DELPHI PROCESS: A METHODOLOGY USED FOR THE ELICITATION OF OPINIONS OF EXPERTS (1968); HAROLD SACKMAN, RAND PAPER NO. R-1283-PR, DELPHI ASSESSMENT: EXPERT OPINION, FORECASTING, AND GROUP PROCESS (1974); THOMAS BROWN, RAND PAPER R-944-ARPA, AN EXPERIMENT IN PROBABILISTIC FORECASTING (1973); HAROLD A. LINSTONE & MURRY TUROFF, THE DELPHI METHOD: TECHNIQUES AND APPLICATIONS (1975); HELMER, *supra* note 11. *See also* T.W. WEAVER, EDUC. POL’Y RES. CTR., REPORT RR-7, DELPHI, A CRITICAL REVIEW (1971); Olaf Helmer & Nicholas Rescher, *On the Epistemology of Inexact Sciences*, 6 MGMT. SCI. 25 (1959); A. Kaplan et al., *The Prediction of Social Technological Events*, 14 PUBLIC OPINION Quarterly 93-110 (1950). Because of subsequent modifications and adaptations, Delphi is sometimes also referred to as the Imen-Delphi procedure. *See* David Passig, *Imen Delphi: A Delphi Variant Procedure for Emergence*, 56 HUMAN ORG.: J. SOCIETY FOR APPLIED ANTHROPOLOGY 53 (1997); *see* Martin Hilbert, *Mapping Out the Transition Toward Information Societies: Social Nature, Growth, and Policies 197-231* (2012) (Ph.D. dissertation, U. of Southern California).

<sup>16</sup> *See* Jon Landeta, *Current Validity of the Delphi Method in Social Sciences*, 73 TECH. FORECASTING & SOCIAL CHANGE 467, 469 (2006) (Delphi “quickly accepted and spread rapidly because it provided valuable solutions to problems inherent in the traditional group opinion based on direct interaction” such as “reduction in the influence of some undesirable psychological effects among the participants (inhibition, dominant personalities, etc.), selective feedback of the relevant information, more extensive consideration than to the repetition, statistical results, flexible methodology and simple execution.”); TED GORDON & OLAF HELMER, RAND PAPER P-2982, REPORT ON A LONG-RANGE FORECASTING STUDY (1964) (assessing long-term science and technology trends, including automation, weapon systems, and population control); Shankar Basu & Roger Schroeder, *Incorporating Judgments in Sales Forecasts: Application of the Delphi Method at American Hoist & Derrick*, 7 INTERFACES 18, 18-27 (1977) (new product sales predicted with error rate of only three to four percent as compared to error rates of roughly twenty percent when predictions done through traditional unstructured forecast methods.).

<sup>17</sup> Although, as reflected in the preceding footnotes, there was a good deal written about Delphi during the 1960s and 1970s, there is much less reference to Delphi in academic literature after 1980. This is not to say, however, that academia has completely forgotten Delphi. *See, e.g.*, MICHAEL ADLER & ERIO ZIGLIO, *Gazing Into the Oracle: The Delphi Method and Its Application to Social Policy and Public Health* (1996); Gene Rowe and George Wright, *The Delphi Technique as a Forecasting Tool: Issues and Analysis*, 15 INT’L J. FORECASTING 353 (1999); Green et al., *supra* note 12; *see generally* Landeta, *supra* note 16.

<sup>18</sup> Outside of law, Delphi appears to continue to have a significant following despite having received less scholarly focus during the past 30 years. *See* Landeta, *supra* note 16, at 469-71.

through web-based technology.<sup>19</sup> This is not to say, however, that Delphi use in government policy making is absent.<sup>20</sup>

The Delphi Method seeks to obtain the judgment of expert groups in a structured methodology for gathering expert opinion.<sup>21</sup> It proceeds from the premise that the quality of the collective expert judgments is far superior to an assembly of individual expert judgments.<sup>22</sup> The essence of Delphi is obtaining expert analysis from a panel of experts regarding a particular issue or problem. It has been defined as a method to structure a group communication process that is effective when allowing a group of individuals as a whole to deal with a complex problem.<sup>23</sup> The technique seeks to achieve a systematic and iterative method that can be used to collect opinions (and a consensus, if possible) from a group of experts. The objective of the Delphi Method is a reliable collective opinion from a group of experts.<sup>24</sup>

Delphi “was conceived as a group technique whose aim was to obtain the most reliable consensus of opinion of a group of experts by means of a series of intensive questionnaires with controlled opinion feedback.”<sup>25</sup> Through a facilitator, the experts are asked to provide analysis that is shared with other experts. The analyses are usually given anonymously so that the group will not be influenced by the relative stature of other panellists or intra-panelist relations or professional rivalries.<sup>26</sup> The goal is also to minimize cognitive error such as

---

<sup>19</sup> See, e.g., MAURIZIO BOLOGNINI, *DEMOCRAZIA ELETTRONICA: METODO DELPHI E POLITICHE PUBBLICHE* (2009); Murray Turoff & Starr Roxanne Hiltz, *Computer-Based Delphi Processes*, in *GAZING INTO THE ORACLE* 56-85 (Michale Adler & Erio Ziglio eds., 1996).

<sup>20</sup> See, e.g., Martin Hilbert, Ian Miles & Julia Othmer, *Foresight Tools for Participative Policy-Making in Inter-Governmental Processes in Developing Countries: Lessons Learned from the eLAC Policy Priorities Delphi*, 76 *TECH. FORECASTING & SOC. CHANGE* 880 (2009) (describing use by Latin American governments for policymaking).

<sup>21</sup> See HELMER, *supra* note 11, at 4-5.

<sup>22</sup> See *id.* at 7; see also Jan A. G. M. Van Dijk, *Delphi Questionnaires Versus Individual Group Interviews*, 37 *TECH. FORECASTING & SOC. CHANGE* 293, 293 (1990).

<sup>23</sup> See HELMER, *THE DELPHI METHOD*, *supra* note 11.

<sup>24</sup> See JON LANDETA, *THE DELPHI METHOD: A TECHNIQUE FOR PREDICTING THE FUTURE* (1999).

<sup>25</sup> Landeta, *supra* note 16, at 468.

<sup>26</sup> See Peter O'Brien, *The Delphi Technique and Educational Planning*, 12 *IRISH J. EDUC.* 69, 69 (1978) (noting anonymity as core part of Delphi process); see also NAT'L SEC. AGENCY, DOC. #3928741, *THE DELPHI TECHNIQUE 1-2* (2011), [http://www.nsa.gov/Portals/70/documents/news-features/declassified-documents/cryptologic-spectrum/The\\_Delphi\\_Technique.pdf](http://www.nsa.gov/Portals/70/documents/news-features/declassified-documents/cryptologic-spectrum/The_Delphi_Technique.pdf) (internal report approved for release Dec. 1, 2011).

bandwagon effects,<sup>27</sup> halo effects,<sup>28</sup> or groupthink<sup>29</sup> that discourages reflection and open criticism, which in turn can lead to analytical error.

Once the panel of experts is assembled, their contributions are usually initially collected through answers to questionnaires, often with provision for open-ended narrative commentary. The facilitator collects the panelist inputs and circulates them, obtaining cross-panelist feedback. Anonymity is usually preserved at this stage but may be relaxed or eliminated later so that panellists can confer in-person as group. Through these iterations, as reported by the facilitator, a group report emerges. The goal is consensus as to action, be it a forecast, the

---

<sup>27</sup> See Linda & Charlie Bloom, *The Bandwagon Effect*, PSYCH. TODAY (Apr. 11, 2017), <http://www.psychologytoday.com/us/blog/stronger-the-broken-places/201708/the-bandwagon-effect>. Bandwagon effects occur when a critical mass of persons taking a view creates momentum in favor of that particular view that makes later-reporting or later-deciding participants more likely to “jump on” the bandwagon rather than to express a dissident view. See *id.* The term is most often used to describe electoral politics in which an early front running candidate gains simply from being the front runner and apparent winner. See *id.*

<sup>28</sup> A halo effect takes place when a person accomplished or admired for a particular body of work is given undue credence or deference when engaging in other endeavors – sometimes wholly unrelated to which the person cannot claim expertise. See *The Halo Effect*, ECONOMIST (Oct. 14, 2009), <http://www.economist.com/news/2009/10/14/the-halo-effect> (defining as “the phenomenon whereby we assume that because people are good at doing A they will be good at doing B, C and D. The phrase was first coined by Edward Thorndike, a psychologist who used it in a study published in 1920 to describe the way that commanding officers rated their soldiers. He found that officers usually judged their men as being either good right across the board or bad. There was little mixing of traits; few people were said to be good in one respect but bad in another.”). See generally PHIL ROSENZWEIG, *THE HALO EFFECT . . . AND THE EIGHT OTHER BUSINESS DELUSIONS THAT DECEIVE MANAGERS* (2007). Common examples abound. Star athletes, actors, and other celebrities are presumed to be affable, kind, smart, knowledgeable about public policy and candidates, etc. even though the public knows only what their respective public relations agents have done on their behalf. The opinions of the famous are taken seriously despite the celebrity’s absence of any formal education, training, or experience in the matter.

<sup>29</sup> See IRVING L. JANIS, *GROUPTHINK: PSYCHOLOGICAL STUDIES OF POLICY DECISIONS AND FIASCOES* 11 (2d ed. 1982) (arguing that groups regress to what might be termed a party line or company position rather than testing assumptions and facts as they would if working without the informal social pressures of group conformity). Janis’s most noted example is the stunningly unsuccessful Bay of Pigs invasion of Cuba by anti-Castro exiles in 1961, which was backed by the U.S. Government. See *id.* at 14-47. The Kennedy Administration planners assumed that with word of the military action, a sizeable portion of Cubans would rise up against the Castro Regime, and assumption that proved stunningly incorrect. See *id.* at 22-26; see also *Groupthink*, PSYCHOLOGY TODAY, <http://www.psychologytoday.com/us/basics/groupthink> (last visited Feb. 1, 2019) (“Groupthink occurs when a group with a particular agenda makes irrational or problematic decisions because its members value harmony and coherence over accurate analysis and critical evaluation”). The U.S. invasion of Iraq in 2003 reflects similar patterns. For example, the U.S. assumed that after Saddam Hussein was deposed, the country would keep running smoothly but as a democracy rather than a dictatorship and purged members of Saddam’s Bath political party from the new post-over through government. But because this also purged the government of expertise, the new regime was poorly managed and prone to corruption and ethnic division as the U.S. also underestimated the enmity between Sunni and Shia Muslim groups, which perpetuated violence long after the official war was deemed over, most ironically in former President George W. Bush’s now-ridiculed “Mission Accomplished” speech rendered aboard an aircraft carrier on May 1, 2003, which preceded the vast bulk of war-related casualties.



root of a problem, or response to a problem. However, “[l]ater applications of the technique have eliminated the restriction of the obligatory search for consensus, so that today [Delphi] might be defined as a social research technique whose aim is to obtain a reliable collective opinion using a group of experts.”<sup>30</sup>

A Delphi inquiry is marked by repeated follow-up, controlled feedback, and an effort to assess group opinion “quantitatively and statistically.”<sup>31</sup> Although Delphi continues to have significant support in policymaking and organizational analysis of problems, it has received more than a little criticism as well, summarized by one Twenty-First Century scholar.

[The Method’s] basic source of information (who is expert, what biases each expert has, etc.) is the use of consensus as a way to approach the truth, the limitation of the interaction involved in written and controlled feedback, the restriction to the possibility of social compensation for individual contribution to the group (the reinforcement and motivation normally provided by the support and social approval of the other expert group members are removed), the impunity conferred by the anonymity with respect to irresponsible actions on the part of the experts, the ease inherent in the methodology of interested manipulation by the person running the study, the difficulty of checking the method’s accuracy and reliability, the time required to carry it out, the effort required on the part of the participants, and the non-consideration of possible inter-relationships between the forecast incidents.

In addition to these methodological weaknesses, the Delphi has also received criticism that is not due to the technique itself but to its deficient application, such as the not very rigorous selection of experts, the lack of explanation concerning its evolution and dropout, questions and problems that are badly formulated, insufficiently analysed results, etc.<sup>32</sup>

Notwithstanding the validity of many of these criticisms, Delphi – at least when done reasonably well -- appears to work as effectively as other methods of inquiry

---

<sup>30</sup> See Landetta, *supra* note 16, at 468 (also describing Delphi as “a method of structuring communication between a group of people who can provide valuable contributions in order to resolve a complex problem.”).

<sup>31</sup> See *id.* at 468-69.

<sup>32</sup> *Id.* at 469 (endnotes omitted); see also Van Dijk, *supra* note 22, at 293-304; Gene Rowe & George Wright, *The Delphi Technique as a Forecasting Tool Issues and Analysis*, 16 INT’L J. FORECASTING 353, 353 (1999); R.C. Erffmeyer & I.M. Lane, *Quality and Acceptance of an Evaluative Task: The Effects of Four Group Decision-Making Formats*, 9 GRP. & ORG. STUDIES 509, 521-525 (1984) (noting studies critical of Delphi Method but finding it provided relatively superior performance for test problem); Wilfred G. Rieger, *Direction in Delphi Developments: Dissertations and Their Quality*, 29 TECH. FORECASTING & SOC. CHANGE 195, 195 (1986); Vaughn E. Huckfeldt & Robert C. Judd, *Issues in Large Scale Delphi Studies*, 7 TECH. FORECASTING & SOC. CHANGE 75, 75 (1974); Gilbert Becker & Donald A. Bakal, *Subject Anonymity and Motivational Distortion in Self Report Data*, 26 J. CLINICAL PSYCH. 206, 209 (1970).

and problem-solving,<sup>33</sup> although its popularity may be hampered because it imposes a potentially time-consuming and expensive layer of logistical and managerial work.<sup>34</sup> Where budgets or time are tight, use of the Delphi Method may involve corner-cutting that result in an unsuccessful experience for the users, who will then be understandably reluctant to use Delphi for future projects, even if the reluctance grows out of their own error.<sup>35</sup>

For that reason, although we advocate greater use of the Delphi Method in law, we do not advocate a Delphi inquiry for every legal problem or every run-of-the-mill dispute. Delphi's greatest potential is as an aid to legislation and rule-making by administrative agencies as well as for executive action with wide-ranging potential impact. Where a legal decision (regardless of the decision-maker), is likely to have impact that is widespread (e.g., a statute), repetitive (e.g., a rule applied to issues that arise frequently), or simply important (e.g., resolution of a billion-dollar dispute), the stakes may merit the searching, sustained analysis envisioned by the Delphi Method.

Individual litigation may often be of insufficient magnitude to support a full-fledged Delphi analysis of the dispute at hand. However, much of adjudication depends on the content of civil litigation codes and rules, which in turn can make an individual case important if it is likely to result in a precedential interpretation of a rule or statute.

Cases of first impression might be particularly apt for Delphi Method analysis. The alternative is that the first case presenting an interpretative issue may be poorly decided in an ad hoc manner without sufficient input from experts. Because the case is one of first impression, a bad decision is likely to have undue influence on the development of law on this topic. However, in cases like these, there will undoubtedly be concerns about whether a potentially precedent-setting ruling is one purely of law, and therefore, to be made by the court without fact-finding. Nonetheless, in apt cases where a court's legal ruling hinges at least in part on an analysis of disputed fact or the overall context of the dispute, the Delphi Method may be an apt tool that enhances the judicial function rather than encroaching upon it.

Although rule makers are not inattentive, neither does one find frequent revision of rules or statutes to keep pace with developments. The process, at least in the United States, has been characterized more by anecdote and reaction to U.S. Supreme Court decisions rather than by any sustained empirical and expert examination of the administration of litigation.<sup>36</sup> Consequently, we see rulemaking

---

<sup>33</sup> See Landetta, *supra* note 16, at 469-74; see also Chia-Chien Hsu & Brian A. Sandford, *The Delphi Technique: Making Sense of Consensus*, 12 PRACTICAL ASSESSMENT, RES. & EVALUATION 1, 1-2 (2007).

<sup>34</sup> See Landetta, *supra* note 16, at 469-70 (discussing Delphi methodology).

<sup>35</sup> See *id.*

<sup>36</sup> This is not to denigrate work done by the Federal Judicial Center and the Administrative Office of the U.S. Courts, as well as rule maker consideration of scholarly studies, all of which takes place to some degree. See, e.g., Emery Lee III & Thomas Willging, *Defining the Problem of Cost in Federal Civil Litigation*, 60 DUKE L.J. 765, 767-71 (2010). Our point is both that there should be more such

as an avenue for greater use of the Delphi Method – and “bottom up” (rather than “top down”) rulemaking generally.<sup>37</sup>

And while it is not reasonable or realistic to think of empanelling a Delphi expert group for most typical litigation, a Delphi approach may be useful for assisting courts in resolving high-stakes or precedent-setting litigation, although the approach may need some degree of streamlining in order to accommodate the need for a relatively prompt decision.

---

empirical work and that the implications of such work should more often be assessed by experts in structured fashion (i.e., the Delphi Method, or something like it) than is currently the case. For example, an empirical study done by the government may permit several alternative inferences. These are now tested, if at all, informally and perhaps slowly through extended dialogue in the scholarly literature – perhaps. If the Delphi Method is more regularly harnessed to assess empirical work and studies of courts, agencies, and statutes in operation, this can improve the analysis.

<sup>37</sup> By “bottom up” rulemaking, we mean rules that emerge from a study or system that in turn prompts rule makers to revise current rules and procedures. For example, an advisory committee on rules may commission a study—e.g., using the Delphi Method, or otherwise gather data on a problem or situation, assess the situation, determine if rules revision is required, and promulgate new or revise rules (or repeal old, outdated, malfunctioning rules). By contrast, “top down” rulemaking takes place when rule makers are more or less directed to revise rules (or leave written rules out of sync with rules in operation) by adjudicator decisions that, if you will, “get ahead” of the rulemaking process by rendering a decision that has the functional effect of changing a rule’s operation without benefit of a formal amendment of the rule. Regarding civil litigation, the United States federal (national) court system has seen both. One example of top-down rulemaking, in 1986, the U.S. Supreme Court decided three summary judgment cases that made summary judgment easier to obtain and equated summary judgment and what was then called “directed verdict.” See *Celotex Corp. v. Catrett*, 477 U.S. 317 (1986); *Anderson v. Liberty Lobby*, 477 U.S. 242 (1986); and *Matsuhita Elec. Indus. Corp. v. Zenith Radio Corp.*, 475 U.S. 574 (1986). A grant of summary judgment had previously required no significant evidence favoring the nonmovant while to preclude directed verdict, the nonmovant must have substantial evidence in its favor. The 1986 “Trilogy” of cases essentially collapsed the two standards, prompting revision of FRCP Rule 50 and changed in its nomenclature form, directed verdict, to “judgment as a matter of law.” See FED. JUD. CTR., *THE ANALYSIS AND DECISION OF SUMMARY JUDGEMENT* 1, 6-7 (1991). Sometime later, Rule 56, the summary judgment rule, was also revised to conform to the case law established by the trilogy and its wake. See BROOKE D. COLEMAN ET AL., *LEARNING CIVIL PROCEDURE* 598-603 (3d ed. 2018); Jeffrey W. Stempel, *A Distorted Mirror: The Supreme Court’s Shimmering View of Summary Judgment, Directed Verdict, and the Value of Adjudication*, 49 OHIO ST. L.J. 95, 185 (1988); D. Michael Risinger, *Another Step in the Counter-Revolution*, 54 BROOK. L. REV. 35, 43 (1988). By contrast, the 1993 Amendment to FRCP Rule 11 was an example of bottom-up revision. A series of court decisions, some by the U.S. Supreme Court, resulted in what many perceived as harsh and unduly formalistic application of Rule 11, which had been revised in 1983 to be more stringent, and in general provided that litigants bringing claims that were not “well grounded in fact” or “warranted by law” were subject to mandatory sanction. See GREGORY P. JOSEPH, *SANCTIONS: RULE 11 AND OTHER POWERS* 63 (4th ed. 2017). Based on negative experiences with the 1983 version of the Rule, the Advisory Committee on the Federal Civil Rules revised Rule 11 to provide greater flexibility regarding sanctions, make it less of a potential fee-shifting rule, and to give more breathing space to law reform claims. See *id.* In addition, law firms were made jointly responsible with their attorneys signing pleadings, motions, and other papers subject to the rules. See *id.*

### III. THE DELPHI METHOD IN OPERATION

The Delphi Method can be used for two main objectives: predictive objectives and opinion collection.<sup>38</sup> Predictive objectives seek to obtain information about future scenarios.<sup>39</sup> As noted above, the most well-known Delphi use is forecasting through the systematic collection of opinions and information on a specific theme.<sup>40</sup> This application is particularly relevant when historic data are lacking as it allows a wide range of interrelated variables to be collected.<sup>41</sup>

#### A. Premises

The following *premises* underly this aspect of the Delphi research methodology.

First, the future is constructed and does not happen simply out of past habit. Rather, the future results from a culmination of interacting forces, many of which are the result of decisions and conduct, whether conscious or unconscious, rather than mere chance.<sup>42</sup>

Second, objective forecasting techniques rely on major limitations that fundamentally derive from insufficient objective information being available in suitable form for analysis. Almost by definition, forecasting must be done without sufficient information.

Third, in certain circumstances, using subjective information is not only justified, but it is desirable for assessing value judgments and expert recommendations. For example, if individuals (e.g., consumers) find certain foods enjoyable even if not nutritionally recommended, this has implications for the efficacy of food regulation. Banning Twinkies might be based on good science, but it also results from sufficiently horrible economics, sociology and politics to preclude such a ban.

Fourth, an individual expert's judgment is always subject to limitations. For example, the expert may have limited available knowledge, which affects the

---

<sup>38</sup> See NORMAN C. DALKEY & DANIEL L. ROURKE, EXPERIMENTAL ASSESSMENT OF DELPHI PROCEDURES WITH GROUP VALUE JUDGMENTS 1 (1971).

<sup>39</sup> See Green et al., *supra* note 12.

<sup>40</sup> See Robert Fildes, Mike Jalland & Doug Wood, *Forecasting in Conditions of Uncertainty*, 11 LONG RANGE PLANNING 29 (1978).

<sup>41</sup> See Uma Gupta & Robert Clarke, *Theory and Applications of the Delphi Technique: A Bibliography (1975-1994)*, 53 TECH. FORECASTING & SOC. CHANGE 185 (1996).

<sup>42</sup> To be sure, part of the future is the result of factors that are at least immediately beyond human control. A lengthy, widespread drought, for example, will place socioeconomic and political pressure on society. Political upheaval often follows food shortages. And although experts agree that humans have had an impact on climate, humans cannot control particular weather. In similar fashion, the unregulated activity of businesses or consumers in a market economy is of course at least partially beyond the control of legal and political decisionmakers. But not entirely. For example, one might not have predicted Amazon or Uber. But the manner in which governments tax online sales or permit livery services (e.g., with or without permit; minimum insurance requirements; taxation regime) affects the outcome of these market innovations.

accuracy of estimations. In extreme cases, experts may be basing assessments on outdated or erroneous information.

Fifth, the quality of the judgment made by the group of experts is better than that of an individual in terms of both the large amount of information available and the positive effects that interaction in a group promotes.

Sixth, judgments made by expert groups also have inherent limitations that affect the quality of the result, which derive from group thought, noise or pressure to reach an agreement as well as the ideological preferences of the expert.

### **B. Assumptions**

In addition, there are the *assumptions* that are the basis of such research (which are normally accepted), such as

- The relevant information that a group of experts accumulates is equal or larger than that of any group member;
- Any incorrect information the group of experts has equals or is larger than that of any group member;
- The number of informal models that the group must construct or estimate will be as large as the broadest preference of a group member;
- Consequently, the number of mistaken informal models that the group can possess and apply to make an estimate is as large as that which any group member may have;
- The median response to a numerical estimation is at least as good as half the individual responses provided by group members;
- An expert who requested a numerical estimation perhaps has only an intuitive and poorly defined probability distribution about this quantity. Consequently, the expert to a certain extent will tend to express the central tendency of this distribution. Therefore, a more accurate estimation can be obtained by aggregating the group members' different individual distributions and selecting the mean (or median) of the resulting distribution as a group response; and
- By means of controlled feedback and by maintaining the experts' anonymity, many troublesome psychological effects can be eliminated from the face-to-face discussion groups without having to the reject advantages of group interaction and deliberation. That said, there is a trade-off. Avoiding face-to-face interaction reduces the risk of group think, bullying, or deference to the more prestigious experts. But in return, the intimate and immediate conversation and synthesis of experts is lost.

### C. Characteristics

After consideration of the premises and assumptions, the following distinctive characteristics of the Delphi Method emerge:

- **Iterative Process:** the experts who intervene in the process must give their opinions on several occasions. Through successive rounds, their estimations tend to converge, and the process ends when opinions are stable;
- **Anonymity:** no participant knows the individual responses of any other group member. The objective is to eliminate the potential causes of inhibition in participants, and the action of dominant individuals;
- **Controlled Feedback:** transmitting the group's general position at each point of the process and identifying significant disagreements that may emerge. In this way, experts know the results reached in the former round when the next round begins. Controlled feedback also allows noise to be eliminated;
- **Statistical Group Response:** when the group is requested to provide a quantitative estimate, the result is normally determined by the median of the individual responses. Although reaching a consensus is sought, it need not be reached to make the process effective. The interquartile range of estimations will indicate the level of consensus reached. Obviously, the dispersion in estimations will also be interesting for the study. In short, pursuing the statistical group response makes all the individual contributions present in the group response. But social pressure for group members to reach a consensus is reduced.

### D. Advantages and Concerns

The main advantages of the Delphi approach lie in its methodology: it allows the strengths of group interaction techniques to become available (many people participate, creativity, etc.) without undue effect by any political, social and personal pressures in the group. The common criticism of direct expert interaction—the emotional component fuelled by views based exclusively on beliefs, feelings and expectations—is minimized<sup>43</sup> (or at least reduced) through careful planning that minimizes the risk.

After defining the problem, participation of experts must be guaranteed. Normally, this process in the Delphi Method consists of sending questionnaires to a group of experts who will make the different estimations. Then, the obtained data will be processed (with a minimum of two rounds).

Experts' estimations are largely dealt with by aggregation to obtain a central distribution tendency measurement (normally the median), taken as a

---

<sup>43</sup> See DALKEY & ROURKE, *supra* note 38; *see also* Gupta & Clarke, *supra* note 41.

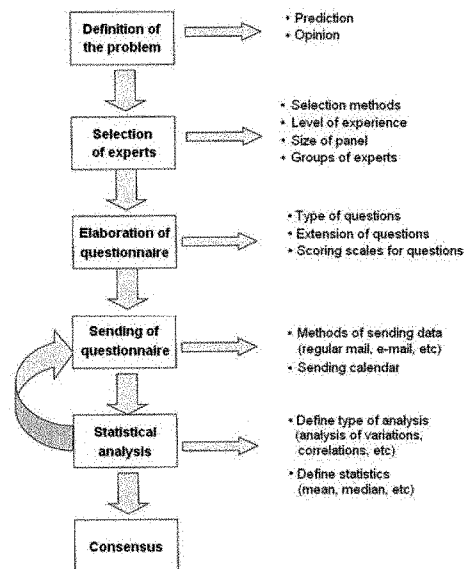
statistical response. The interquartile range is also established as a dispersion measure of the estimations. Then, questionnaires are sent to each expert and include, apart from the information of interest, their former individual responses as the median and the interquartile range of the group as a whole per estimation.

Based on this information, subsequent questionnaires request each expert to review, if necessary, his/her former estimations. If the estimations did not fall within the interquartile range, each expert will be asked to justify the individual opinion with a view to isolating and addressing the non-conforming estimations before arriving at consensus.

This iterative process ends when all the estimations are stable; that is, when the median hardly oscillates, and the interquartile range no longer narrows (with at least two rounds). From this point, all that is left to be done is to take the last round as a group response and to prepare the corresponding report.

The following graph schematically illustrates the Delphi Method stages:

**Figure 1: The Delphi Method Phases<sup>44</sup>**



### E. The Delphi Research Frame: Delimiting the Problems to be Studied

The Delphi Method is applied within an existing frame (e.g., law, business, education policy). Consequently, use of Delphi to assess issues of law and legal policy proceed within the framework of traditional legal methodologies. As we

<sup>44</sup> The diagram above is based on Cristina Chaminade, Pedro Luis Sánchez, & Carmen Gloria Escobar, *En Busca de una Teoría Sobre la Medición y Gestión de los Los Intangibles en la Empresa: Una Aproximación Metodológica*, 45 EKONOMIAZ: REVISTA VASCA DE ECONOMÍA, 188-213. (1999).

have seen, Delphi is a method that employs questionnaires. But it is not easy to focus questions on the relevant and conflictive points of a given matter. Preparing suitable questions alone is a complex task. Investigators must establish a research frame, define the problem, draft questions, and then obtain sufficiently detailed and thorough responses to yield meaningful injections of expertise.

A legal investigation that does not involve such prior work can hardly offer the expected rigor. Experience shows that any legal investigation includes numerous aspects on which there will not be uniform agreement. Therefore, questions must center on the more controversial points. An almost infinite questionnaire should be avoided because it can negatively affect the investigation by discouraging experts from participating.

The Delphi Method does not substitute for traditional legal investigation techniques but is an additional method. Its value lies in complementing them when well-established methods cannot offer a foreseeable solution beyond the author's authoritative opinion of the work or that offered by a small group of jurists, which prevents analysing the probability score in its results.

## **F. Goals of the Delphi Method**

As noted, the Delphi Method can take a predictive or prospective purpose. For the former, it is a matter of requesting expert opinion on matters that have still not taken place, and of obtaining an opinion of what will take place from what has happened. In the latter, the aim is to obtain an opinion about what has happened or is happening right now. Regarding law, this requires consideration of several factors.

### **1. Predictive Orientation**

In the social sciences, the Delphi Method has enjoyed support in part because it is suitable for dealing with problems for which no abundant empirical evidence exists. This entails application to emerging issues that lack substantial empirical study.

If we shift this premise to the legal domain, we believe that the Delphi Method should have a special meaning in investigations into recently implemented legal reforms. Here the traditional approach consists in the researcher making a subjective prediction on the rule being applied. The traditional approach is thus limited in that it cannot do anything else in the absence of additional information such as court rulings or doctrinal writings.

Conversely, the Delphi Method has the advantage of offering a group opinion. Consequently, the predictability of the interpretation of the law, which is in force when the application of this technique begins, may improve. Subjectivism is at least minimized if not eliminated no matter how much authority the author may have, which helps obtain a group opinion because a variety of experts participate and interact in the process.



## 2. A Prospective Orientation

The Delphi Method in the legal domain may go beyond predictive purposes. Indeed, it can also be applied to analyze what currently takes place regarding statutes, regulations, or court decisions. Delphi, as opposed to traditional methods, offers many advantages: reaching opinions supported by a larger number of experts; including research results in practice rather than relying too heavily on theory alone; having a more rigorous tendency to adhere to widely supported interpretations; and including several disciplines to seek a group consensus. Consequently, using the Delphi Method to obtain an opinion about legal issues offers the prospect of improvement in legal analysis.

The Delphi Method thus may be a valuable technique for assessing doctrinal disputes among judicial decisions. The systematic injection of expert opinion widens the spectrum of analysis. When the consultation spectrum widens, the Delphi Method can assess the division and perhaps achieve a result enjoying wider support, even if consensus proves illusive.

Similarly, a Delphi analysis can establish the degree to which consensus is lacking and whether consensus is even attainable. For example, in some areas of law, there may be sufficient division, even among experts, that the only practical solution is to tolerate case-by-case, jurisdiction-by-jurisdiction differences.

The Delphi Method does not hinder the application of doctrinal work or theoretical jurisprudence that predates the collective Delphi analysis. It merely provides another means of assessing existing legal doctrine as well as a means of identifying the need for additional research and study.

## 3. Questions and the Questionnaire

As in any other studies that resort to questionnaires, the wording of the questions is a fundamentally important step. Bias that could affect the end result must be avoided. Questions have to be clearly and directly considered. It is important to ensure that the inquiry is understood by the experts and avoids conditioned responses from the selected experts. The questionnaire should be worded to avoid questions that require extensive interpretation of responses. Conclusions drawn should leave as little margin as possible for subjectivism and should offer clear and objective results.

Beginning the questionnaire with open questions is recommended as it enables the rest of the study to be linked to the provided responses, and thus reduces researcher bias that may affect the phrasing of more pointed questions. But questions in the Delphi process cannot be fully open-ended throughout the process as this would make subsequent statistical analysis difficult if not impossible.<sup>45</sup>

---

<sup>45</sup> However, where the panel of experts used in Delphi analysis is small, it may be possible to proceed with a greater portion of open-ended questions. Where a sample size is small, statistical analysis of the responses is not as valuable as the substantive content of the responses.

Consequently, the types of questions used in Delphi analysis fall into two main groups:

**A. Hierarchies, Evaluations, or Comparisons of Items.**

- 1) *Hierarchy*: assigning relative positions to each item (ordering).
- 2) *Evaluation*: giving points according to a defined scale. If there are many items, it is easier than hierarchy and also transmits more information.
- 3) *Comparison*: a pairwise comparison of items according to the defined criterion.

**B. Specific Quantitative Estimations.**

- 1) *Sporadic Estimations*: They can be both absolute and probabilistic. The former are employed to estimate the future properties of a variable, while the latter are used to estimate the probability of a fact occurring.
- 2) *Non-sporadic Estimations*: They normally appear as confidence intervals on the values that a variable takes at a given confidence level.

Thus, in order to minimize lack of open questions and to obtain the maximum complexity of experts' knowledge, interviews are first conducted with several people in order to seek potentially typical responses. This procedure allowed a first data analysis of the responses obtained in the interviews.

Experience evidences the utility of pretesting to avoid potential questionnaire design defects. For this pre-test, people who belong to the various expert profiles are sought to guarantee that questions are suitable and understood. The benefits that the pre-test offers are evident as they allow questions to be devised that are better understood, and the scale of a given question can also be changed, which significantly improves the subsequent data analysis.

#### **4. Selecting Experts: Profiles; Knowledge; Geography; and Demographics**

Experts represent the core of the Delphi Method as they are responsible for estimations. Consequently, one critical difficulty is to define the *de expert* condition. *A priori*, the characteristics that define it will be their knowledge about the study object and its predictive capacity (in the order of former estimations). Too restrictive a concept of expert qualifications makes it more difficult it will be to conduct a Delphi study.

In recent decades, as Delphi methodology has been applied to ever increasing heterogeneous fields, the definition of an expert comes closer to the *collaborator* concept. Under current Delphi practice, an expert is defined as anyone whose situation and resources allows him/her to positively contribute to achieve the purposes set out in the Delphi study.<sup>46</sup>

---

<sup>46</sup> This is our distillation and characterization of the literature on Delphi, which is surprisingly light concerning expert qualifications and the manner of expert selection. See LINSTONE & TUROFF, *supra* note 11 (no express discussion of criteria for expertise or manner of expert selection in entire book);

Along these lines, the definition of expert has been described as “anyone who provides any relevant input.”<sup>47</sup> Consequently, various roles can be distinguished according to an expert’s role in a Delphi study. These include:

- **Specialists:** they combine knowledge, predictive capacity, experience and objectivity. They represent the classic specialist concept;
- **Affected Persons:** they know the study object but are not distinguished as having more knowledge than the mean. Their interest lies in them being involved in the study object and the Delphi results will be applied to them; and
- **Facilitators:** they are people who do not belong to either of the two above categories, but it is their task to explain, stimulate or [organize].<sup>48</sup>

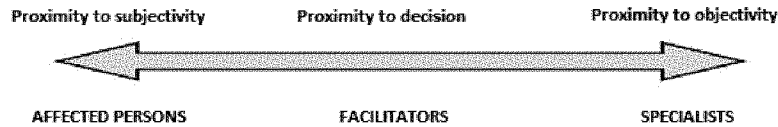
Obviously, these persons in these three expert categories may vary in the degree to which they take more objective or subjective positions regarding the issue under study. This contingency may make some expert input more apt than others in terms of the issue under study. There is no optimum rule for all cases).

---

Part IV concerning experts focuses on group dynamics but does not specifically address criteria for being denominated an expert); Landeta, *supra* note 16, at 468 (referring to experts but saying nothing regarding manner of determining a qualified expert). *See also* LANDETA, THE DELPHI METHOD, *supra* note 24 (suggesting this view of expert qualification in the book generally). *See also* LINSTONE & TUROFF, *supra* note 11, at 566 (discussing “Illusory Expertise” as one of eight pitfalls that may impair Delphi analysis but not discussing qualification of experts per se). For some studies, those applying Delphi are apparently willing to trade interest for expertise. *See, e.g.*, Kristen Dufresne, *The Delphi Technique*, STUDENTS4BESTEVIDENCE.COM (Nov. 15, 2017) (available at <http://www.students4bestevidence.net/author/kristen-dufresnethe-delphi-technique/>) (“Research has indicated that participant subject matter knowledge (i.e., being an “expert”) may not have a substantial impact on study results, so it might be best to choose participants who have some understanding of the topic and an interest in the outcome of the study to limit attrition and encourage thoughtful responses to the surveys” and “[o]ften, participants are selected via non-probability sampling techniques (either purposive sampling or criterion sampling), to save resources and ensure appropriate participants are selected). Notwithstanding this, our assessment of the Delphia literature is that it implicitly endorses mainstream criteria for expertise such as formal education, substantial experience in a field, and academic executive positions.

<sup>47</sup> *See* Juri Pill, *The Delphi Method: Substance, Context, A Critique and an Annotated Bibliography*, 5 SOCIO-ECONOMIC PLANNING SCIENCES 57, 58 (1971).

<sup>48</sup> *See* D. SAM SCHEELE, THE DELPHI METHOD: TECHNIQUES AND APPLICATIONS 35, 65 (Harold A. Linstone & Murray Turoff eds., 1975).

**Figure 2: Types of Experts and Proximity to Objectivity<sup>49</sup>**

In this way, given the composition of the expert group, beyond its specific application to Law, the following may occur:

- a) **Preponderance of Specialists:** when information for decision-making is needed about the evolution of external variables to the group and decision maker;
- b) **A Majority of Affected People:** when a conflict of interest exists, or when the method is intended to be used as a communication or learning instrument; and
- c) **Clear Presence of Facilitators:** required when facts, values or relations are not clear

Hence the works conducted with it stand out as they rely on the preponderance of “specialist” experts, which evidences a systematic objectivity search approach. The selection of specialists has been traditionally made by reputation and opportunity criteria.<sup>50</sup>

In the legal domain, sought specialists tend to belong to two typical profile types: academic and professional. Other groups may emerge when focusing on a given investigation, but the above two groups are definitely the most usual ones.

When we mention *academics*, we refer to those university teachers, professors or doctors who have investigated the given field which the questionnaire is about. Application of the academic reputation criterion is done by inviting experts with outstanding publications in the given study fields.

We understand *professionals* to be those people whose professional activity centers on practically applying the rulings that are the research object. Here we clearly include those lawyers who specialize in this problem, but other professionals can also be included. The two notes that must, at least, apply to the selected people are solid training in the matter and being constantly in contact with the practical application of the law in question. In this segment, reputation derives from institutional responsibilities (professional associations, etc.), relevance in mass media, etc.

<sup>49</sup> The diagram is based on LANDETA, THE DELPHI METHOD, *supra* note 24, fig. 3.2.

<sup>50</sup> This history prompts the question of whether the Delphi Method could use an injection of its own principles regarding selection of experts. A prominent reputation may be well-deserved or inflated. Perhaps selection of experts should be according to a checklist of objective criteria such as published writings, citations by scholars, citations by courts, retention as a consultant, or other criteria less amorphous than reputation alone.

Regarding the decision makers in legal investigations, it is worth including among experts those legal parties with the capacity to apply the law through binding rulings. Here we obviously refer to judges and magistrates with the objective competence of having the capacity to sentence when the law is applied, but they are not the only ones. This category can also include civil servants who are not judges, but who have the capacity to sentence binding rulings. Likewise, we also include arbitrators on the list of an arbitral institution. Basically, while professionals remain in contact with the law, applicators overcome any controversial matters of the law in a *binding* manner.

In addition, the group may include persons who are not experts per se but have information and perspectives to contribute because they are affected or will be affected by the issue under study. For example, in a study of insurance products, laypersons such as policyholders may be included. So, too, may persons with experience in the field such as claim adjusters, even though these persons are not ordinarily thought of as experts in the manner of a scholar or top management.

As indicated, the response rate of dispatched invitations tends to be low, but it is important to stress its wide variability according to various expert profiles; e.g., the response rate tends to be higher for the academic group than for the professional group and the group engaged in the issue. An attempt to overcome this phenomenon is made by sending more invitations to experts who belong to the segments that foreseeably provide a lower response rate, and a search for specific prescribers can be made.

In order to identify existing differences among different groups, studies should attempt to strike a certain balance. That is, researchers must be very careful with the geographical area of the group of experts as outcomes could be conditioned by a local law that is not generalized to the whole area where the law is applied. However, different response rates per sector imply a heterogeneous presence of profiles. In principle, this fact does not affect the quality of the group response if a minimum number of experts has been exceeded; greater participation barely affects a mistake made but makes the segmented analysis difficult (according to territories, professional profiles, etc.).

As a result, and according to our practical experience, the importance of the opportunity criterion has been evidenced when forming a group of experts because access to them determines the response rate to a great extent.

Finally, it is worth referring to experts' degree of experience, which was heterogeneous in all the research done. In order to bear this fact in mind, a self-assessment about the various matters relating to the study of each expert's knowledge is made. It appears that an expert's self-assessment is not significantly worse than an assessment by a third party.<sup>51</sup>

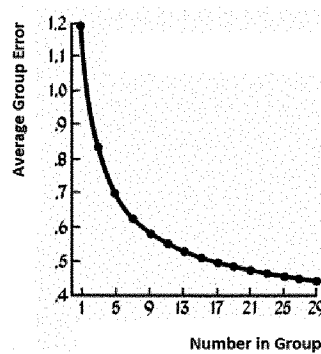
---

<sup>51</sup> See Landeta, *supra* note 16.

### 5. The Size and Deployment of Expert Groups

There is no way of exactly determining the optimum number of experts that form a group. This question has been approached from the correlation between the number of individuals who form the group and its exactness in estimations. Likewise, empirical experiments have demonstrated that as the number of group members grows, the error in the final forecast (measured as the median moving away from the final value) exponentially diminishes. The figure below illustrates the relation between the number of experts and the error in forecasting.

Figure 3: Effect of Group Size on Error in Final Forecast<sup>52</sup>



If we bear this function in mind, a minimum of seven experts is generally recommended to ensure an acceptable mean group error. If we look at the cost-benefit analysis, more than thirty experts would not be advisable since the efforts made to include a large volume of estimations do not necessarily mean improved forecasting.

In the analyzed studies, the expert group size has not usually been a problem because groups normally involve more than fifteen experts. It is noteworthy that the response rate to invitations is low as the number of invitations sent is quite large (between several dozens and hundreds depending on each case). Low response rates for questionnaires can be minimized by ensuring that experts are fully informed about the study and that reminders are issued.<sup>53</sup> Resorting to prescriptors who contact experts personally significantly helps increase the response rate.<sup>54</sup>

<sup>52</sup> See Norman Dalkey & Daniel Rourke, *Experimental Assessment of Delphi Procedures with Group Value Judgments* (Advanced Res. Projects Agency, Working Paper No. R-612-ARPA, 1971); NORMAN C. DALKEY, RAND No. 5888-PR, *THE DELPHI METHOD: AN EXPERIMENTAL STUDY OF GROUP OPINION*, at 11 (1969).

<sup>53</sup> See COLIN ROBSON & KIERAN MCCARTAN, *REAL WORLD RESEARCH* (2016); LOUIS COHEN & LAWRENCE MANION, *RESEARCH METHODS IN EDUCATION* (4th ed. 1994).

<sup>54</sup> See Chia-Chien Hsu & Brian Sandford, *Minimizing Non-Response in The Delphi Process: How to Respond to Non-Response*, 12 *PRACTICAL ASSESSMENT, RES. & EVAL.* 1 (2007); see also Elizabeth

## 6. Expert Motivation, Time, and Available Resources

The quality of questionnaire-based works largely depends on the degree of experts' motivation. Factors like questionnaire length, number of rounds or the ease with which to give responses certainly determine the response rate and the quality of responses.

If we consider that experts receive very few incentives (they are usually offered only the final research report), it is very important that the questionnaire is concise and easy to complete. Attempts have been made for completed questionnaires to contain twenty to thirty questions, and they have been designed by using tables to help responses to be quickly made using crosses or scores.

Inquiry to experts then enters a second round in which there is highlighting of the responses that were not reached by consensus. Experts are prompted to reconsider them (by facilitating responses in both time and type terms). A strength of the Delphi Method is its iterative process, which allows experts to provide their opinions in successive rounds, which allows for refinement of those opinions.

Likewise, the number of study rounds also has a direct effect on questionnaire response rates. Consequently, participant burden can be a problem, particularly when the number of rounds exceeds four.<sup>55</sup> This suggests perhaps limiting a study to only two rounds for maximum benefit. Although this is only the minimum set by the method, it reduces the problem of declining response rates the creation of a possibly unrepresentative sample of experts. For example, experts of a particular area may be more or less tenacious in continuing to remain involved through multiple rounds.

A design that includes only two rounds entails one main drawback: it is not possible to suitably process the "new" responses provided by experts (not known by the rest in the first round). This means that their statistical analysis is not feasible, although responses are employed at the descriptive level. Yet in cost-benefit terms, our experience indicates that a process with two rounds is the most efficient design for this type of research.

## 7. Concept and Consensus Determination

The Delphi Method determines the median ( $m$ ) as the central tendency measurement of the expert group response; that is, the median is the response of the group included in the study. The interquartile ( $k$ ) range is also calculated to measure sample dispersion, which is inversely proportional to the group consensus (e.g., the wider the range, the more difficult the consensus). As complementary indicators, the following are also employed: the arithmetic mean ( $\mu$ ), the mode

---

Gargon, et al., *Higher Number of Items Associated With Significantly Lower Response Rates in COS, Delphi Surveys*, 108 J. CLINICAL EPIDEMIOLOGY 110 (2019).

<sup>55</sup> See Hsu & Sandford, *supra* note 53; Gargon et al., *supra* note 53.

(*Md*) and standard deviation ( $\sigma$ ). These values are particularly useful for establishing a relative order between items with the same median.

In this context, unanimity is accomplished when  $k = 0$ , or when the relative frequency of the response is above or equals eighty percent. One of the main difficulties of this method lies in estimating the level of acceptable convergence (consensus) among experts when this circumstance does not occur. As a result, the wider the interquartile range permitted in that understood as a consensus, the easier it will be to reach it, but the harder it will be to reduce the dispersion of the responses in later rounds. Hence determining a 'balanced' interquartile range is one of the most sensitive points of applying this technique since determining a consensus entails certain subjectivity and a discretionary nature by the coordinator group.

It should be noted that achieving a useable interquartile range appears to depend on the number of possible responses (from the *scale*). It is not feasible to expect the same degree of consensus for two options as for ten. The interquartile range should have an adequate number of dichotomic questions and should widen according to the number of possible responses. In practice, the following values were taken in research work<sup>56</sup> to establish the consensus:

---

<sup>56</sup> See Elies Seguí-Mas, Juan Bataller-Grau & Javier Vercher Moll, *Los Métodos de Resolución de Conflictos en el Contrato de Seguro: Un Estudio Delphi en el Marco de la Unión Europea*, UN DERECHO DEL SEGURO MÁS SOCIAL Y TRANSPARENTE, 297, 297-322 (2016); Elies Seguí-Mas and Juan Bataller-Grau, *La Protección del Asegurado en la Unión Europea: Una Aproximación Desde la Metodología Delphi*, In III Congreso Nacional de Ordenación, Solvencia y Supervisión en seguros privados y II Congreso Internacional de Derecho de Seguros: Valladolid, 17 y 18 de noviembre de 2016, PSYCOM DISTRIBUCIONES EDITORIALES 157, 157-174 (2015); Elies Seguí-Mas, *Factores Clave en la Protección del Asegurado: Estudio Empírico a Través de la Consulta a Expertos*, LA REFORMA DEL DERECHO DEL SEGURO 59, 59-71 (2014); Thomson Reuters-Civitas; Juan Bataller-Grau & Elies Seguí-Mas, *La Protección del Cliente: Un Estudio Empírico*, LA PROTECCIÓN DEL CLIENTE EN EL MERCADO ASEGURADOR 49, 49-106 (2013); Sergio Marí Vidal, María del Mar Marín Sánchez, Elies Seguí-Mas & Marta Michael-Zamorano, *La Distribución Como Factor Determinante en la Protección del Asegurado*, ESTUDIOS SOBRE MEDIACIÓN DE SEGUROS PRIVADOS 119, 119-160 (2014); Asociación Interprofesional para el Desarrollo Agrario; Elies Seguí-Mas & J. Server Ricardo, *Análisis Cualitativo de los Determinantes en la Predicción del Fracaso Empresarial en Cooperativas Agroalimentarias*, 110 INFORMACIÓN TÉCNICA ECONÓMICA AGRARIA 300, 300-320 (2009); Elies Seguí-Mas & J. Server Ricardo, *Studying the Financial Resources for Agri-Food Industry and Rural Development: Description of Human Capital in Credit Unions Through Delphi Analysis* 34 Interciencia 10 (2010); Elies Seguí-Mas & J. Server Ricardo, *El Capital Relacional de las Cooperativas de Crédito en España: Un Estudio Cualitativo de Sus Intangibles Sociales Mediante el Análisis Delphi*, 101 REVESCO. REVISTA DE ESTUDIOS COOPERATIVOS 107-131 (2010); Elies Seguí-Mas & J. Server Ricardo, *Las Cooperativas de Crédito y Su Entorno en el Contexto de la Crisis Bancaria: Análisis de Su Capital Relacional Como Base Desde la que Explotar Oportunidades*, CIRIEC-ESPAÑA, REVISTA DE ECONOMÍA PÚBLICA, SOCIAL Y COOPERATIVA 68 (2010); Elies Seguí-Mas & J. Server Ricardo, *Caracterización del Business Capital de las Cooperativas de Crédito a Través del Análisis Delphi*, 103 REVISTA DE ESTUDIOS COOPERATIVOS 101, 101-122 (2010).



- a) In questions with 2 or 3 options:  $k = 0$
- b) In questions with 4 or 5 options:  $k \leq 1$
- c) In questions with more than 6 options:  $k \leq 2$

For the questions with sporadic estimations,  $k$  was determined in a personalised fashion. In these questions, we opted mainly to create intervals and to apply a  $k$  according to the number of options, which is a similar criterion to that previously adopted.

It is worth pointing out that the stability criterion is the equivalent to the consensus criterion in the first round; that is, if a consensus is reached, the result will be considered stable. In the second round however, the stability criterion is independent of the consensus criterion. Stability in the group's response is understood as this response changing in the short term (irrespective of a consensus having been reached or not) being quite unlikely. To evaluate it, a relative interquartile range ( $r$ ) was used for each response.

By defining the relative interquartile range as the interquartile range divided between the median, its variation ( $Vr$ ) will be the equivalent to the difference between the relative interquartile ranges of the two successive rounds; that is,  $Vr = r_j - r_i$ . Thus, when the relative interquartile range falls between -0.25 and 0.25, a satisfactory level of group response stability is obtained, which will probably not change in successive rounds. Irrespective of whether consensus is reached, the response will be considered stable.

## **8. Managing feedback**

Feedback to the group of experts in the second round (and any later ones) is fundamental as its management can prove to be a strong bias that can either facilitate or frustrate statistical consensus in the group. Consequently, it is important for the coordinating group's feedback interaction that the group follow clear and stable criteria throughout the investigation.

The main difficulty lies in using open responses because, in order to facilitate subsequent analysis, these less easily coded responses must be characterized and grouped. This means that information may be lost during this process. Likewise, feedback management must guarantee that none of the contributions made is ignored.

## **IV. THE ADVANTAGES OF THE DELPHI METHOD**

The Delphi Method described above can add value as either a substitute for or in addition to traditional methods of legal analysis. Although we see most potential for its use in the broader realms of legislation and rule-making, Delphi can provide valuable analysis for adjudication as well as for non-litigated individualized questions such as categorization for purposes of the application of

rules and statues (e.g., should a given high-crime neighborhood be subject to curfews, particularized policing, etc.). We offer our suggestions generally without seeking to offer highly specific examples of the application of the Delphi Method to legal problems, issues, or disputes.

### A. Broader Majorities and Improved Probability and Predictability

Good scientific work has to thoroughly collect all the doctrinal articles which have previously dealt with the topic to be studied. It is normally considered sufficient to compare the number of authors who are included in footnotes to verify this statement. Reliability is entrusted to the researcher's authority criterion without it having to be fully backed by other authors.

Conversely, the Delphi Method allows the opinion of many experts to be included in the work who, without having previously taken written positions, have an opinion as specialists on a given matter.

We believe that if a group consensus is reached in a work that resorts to using the described technique to make an interpretation, its result is much more likely to be accurate or to enjoy extended support than if it is the mere opinion of one author. The Delphi Method is based in part on the view that group analysis is generally more accurate than individual opinion, provided that both rest upon the same facts and follow a similar mode of analysis.<sup>57</sup>

In general, group opinion—if emanating from a group of experts with sufficient academic credentials—works to suppress the subjective biases that may afflict a single expert. The group opinion normally contains richer and more accurate information. Adding the interactions conferred by successive rounds of consultation works to obtain an even more reliable result. Although group consensus is not possible in all instances, it should serve as a relatively accurate barometer of the topic studied and the correct direction for law or policy addressing the problem.

In short, the Delphi Method contemplates an improvement over traditional methodologies and should provide more accurate assessments and better proposed solutions. Group analysis confers more robustness to the interpretation obtained by diminishing biases of an individualised opinion. The group possesses better quality information and reduces the risk of error. To this advantage we add the iterative element of successive rounds of inquiry and consultation that help form group opinion.

---

<sup>57</sup> There are, of course, exceptions to this general rule. For example, when Nicolaus Copernicus advanced the now-confirmed theory that the sun is the center of the solar system (and that earth revolves around the sun and not vice-versa), he was alone in this view. Any group of scientists assembled for a Delphi analysis would have subscribed to the earth-centric view. However, Copernicus was not standing as a lone minority view within the regime of normal science. Copernicus was following a completely different construct and ushering in a new paradigm. *See generally* THOMAS KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* (2d ed. 1974). Consequently, the heroic stories of trailblazing iconoclasts who saw the errors of a given science or social paradigm do not vitiate the value of obtaining group assessment via the Delphi approach or similar methods.

### **B. Fusing Expert Opinion with Law Practice and Policymaking**

One concern is that the Delphi method may approach real world problems too theoretically and thus have limited practical value. Although this is a risk, we think it a manageable one. First, Delphi offers a much more rigorous analysis than traditional method of casually enlisting experts in decision-making.

Second, if the Delphi Method is applied correctly, the panel of experts will include a significant proportion of persons with real world experience. In addition to panellists actively working in the field, many panellists from academia will have had substantial line experience prior to becoming academics. Many will maintain ties to the outside world through consulting and conferences as well.

Third, the perception that panelists from academia are woolly-headed, ivory tower types clueless about the outside world is simply a stereotype that is as baseless as it is offensive. If an observer suggested that CEOs of large corporations were ignorant of history and poorly read, society would quickly recognize this as a stereotype-turned-slur. Considering professors or researchers to lack sensitivity to practical concerns is a similar misconception.

Much depends on the selection of the panel, of course. But a sufficiently diverse panel of true experts (as contrasted to ideologues employed by left-leaning or right-leaning think tanks) should be sufficiently sensitive to practical concerns as well as avoiding individualized bias or an assessment hidebound by a single expert's own personal experiences.

Incorporating practice into study can infuse the Delphi analysis with an undeniable advantage. A well-designed questionnaire can also elicit valuable information and analysis even when administered to a respondent/expert who is otherwise unfamiliar with academic conventions.

The Delphi Method does not deny the value of jurisprudential learning or legal theory, just as it does not ignore empirical data and issues of practical application. But as compared to conventional scholarship regarding legal theory or legal doctrine, the Delphi Method can provide a much broader response.

Delphi attempts a wider inquiry than simple reflection about prevailing theory, which is always conditioned by the background of the contemplated case and is limited by its specific solution whose sole purpose is to reach a conclusion, be it on the form of a statute or rule, a classification, or a ruling in a lawsuit. Conversely, the questionnaire provides complete freedom to deal with those questions not solved by jurisprudence or to ask these questions more thoroughly and without regard to the specific circumstances that led to the legal ruling. No matter how broad jurisprudence is, it will always be conditioned and limited. A questionnaire permits us more autonomy when considering the object of the study.

Moreover, the Delphi Method allows us to question the reliability of the decision made. If a group opinion that emerges is politically controversial or runs counter to conventional wisdom, it nonetheless has better prospects for being acknowledged and accepted than individual work because it results from a

consensus of expert opinion. This may even convey enough credibility to overcome adverse erroneous judicial precedent.

The use of a group of experts from different backgrounds allows diverging opinions to be verified depending on their origin. In this way, not having reached a consensus in the group can be enriched by the dissimilar results obtained from the various groups that make it up. It is not surprising that academics could take a view that is far removed from the consolidated opinion of professionals or decision makers, which thus praises the virtues of this method.

### **C. Objectifying Positions in the Interpreting of Law**

One of the problems faced by the doctrine when preparing a study using conventional methodologies is to reliably reproduce analytic results. Certainly, a doctrinal work must include what has been said to date by other authors, which spells a dilemma that is not easy to solve.

The first possibility is to summarize previous works. This provides readers with a systematic, reduced vision of the debate, which normally leads to clarity and concision. The two or three main interpretations of the law are offered to readers, along with their assumed followers, at the bottom of the page. Preparing work in this way entails losing the subtle nuances that each author has made in the author's work, which implies the risk of losing the desired rigor.

The second possibility is to attempt to collect each opinion with all its connotations, which means conducting less systematic work with a poor synthesis where its author offers a different interpretation to all the others. Opinions will be provided without further ado, and each interpretation with a single defender because it will be impossible to redirect subtle nuances to converging positions. Striking a balance is ideal, but this tends to be elusive.

Conversely the Delphi Method, which includes each expert in each interpretation, is more objective and rigorous. Here the risk lies in wording questions that clearly reflect possible subtle nuances, which moves away from bias to guided research. Unlike traditional methodologies, it is no longer necessary to interpret each author's opinion to see if it can be included, or not, in any predefined hypothesis. Rigor is provided because affiliation is direct.

Consequently, the Delphi method can avoid the risk of re-interpreting what previous works have stated to be able to classify them into some predefined interpretation. Each expert involved in the Delphi Method is uniquely and exclusively evaluated by the question set out, and all the experts are evaluated by the same questions, which minimizes the researcher's discretion when it comes to classifying the responses made by experts.

### **D. Controlled Feedback Enabling Additional Reflection by Experts**

The characterizing feature of the Delphi Method is that it asks experts questions in several rounds in order to reach a group consensus. The second round

is not merely a reiteration of the questionnaire. This is because it centres on those questions in which expert opinions differ, and it does so in order to evidence whether a group consensus can lie behind a last reflection or if this consensus does not exist, which would also be a valid result.

This mechanism encourages that, in case discrepancies that arise, such precise discrepancies are presented to the experts. From this point, experts are able to reiterate the discrepancy or reflect and amend their former position. This is all done by maintaining results anonymous to avoid any element of an expert's authority or his/her affinity to others, which could influence an expert's decision.

Basically, a reflection process with the preliminary results emerges, and is completely unlike traditional methods, which cannot request authors voicing their opinion in writing if, since then, they have amended or added a subtle difference to an interpretation made in a former work. It is necessary to consider that there are times when a written work stands far away in time from the citation included in a new work, which means some reforms may have been made in collateral matters or relevant jurisprudence might have emerged. This would mean that asking whether the author still maintains his/her same thesis is advisable. The Delphi Method overcomes these problems through the methodology in which there is not only immediacy in all the responses, but also a last reflection to confer the group opinion greater robustness.

### **E. Casting a Wide Net of Expertise**

One last advantage that the Delphi Method offers is the possibility that the group of experts is made up of experts of various origins. By origin, we do not refer only to the classification already provided regarding differences of opinion among specialists, decision makers and affected people. We now refer to the expert characteristics which originates from particular expertise in specific disciplines.

Legislators frequently employ non-specific legal concepts, which are gradually included according to their jurisprudential inclination. On occasion however, these non-specific legal concepts are also a summons to include other disciplines in making legal assessments. Obvious examples are economics, medicine, chemistry (e.g., for environmental regulation), geology (e.g., for land use regulation), and perhaps even something seemingly far afield such as urban design (e.g., as a means of maximizing crime reduction with minimum intrusion on the public's freedom of movement).

Confining expertise only to lawyers will often provide poorer interpretation. A jurist is unlikely be an expert in these other disciplines. But moving too far away from the legal perspective also entails evident risks. The Delphi Method appears to be a superior technique in that it can, in organized fashion, incorporate perspectives from other disciplines without becoming unmoored from law. It facilitates integration of expertise.

One example is provided by Spanish legislation protecting business discretion. The legislator estimates that if a decision made by the administrators of

a capital society (or directors of a corporation) was made in good faith, with no personal financial interest in the assumed decision, with sufficient information and *according to a suitable decision-making procedure*, then the administrators/directors are shielded from liability.

But this prompts the question of whether a jurist possesses the elements to determine whether a decision has been made based on “adequate” information and procedures. It seems worthwhile for experts who are familiar with the business world—administrators, auditors, consultants, etc.—to participate in the study to shed some light onto what must be interpreted by an “adequate decision-making procedure”.

In short, the Delphi Method enables the inclusion of different professional profiles in the group of experts, which doubtlessly enriches research results, despite the conclusion of showing some divergence among the opinions of all the groups that make up the group of experts.

#### **V. CONCLUSION: THE DELPHI METHOD AS A BRIDGE TO MORE SCIENTIFIC LAW**

Like any mode of analysis, the Delphi Method is open to criticism. While we have anticipated and attempted to respond to much of the criticism, we acknowledge that Delphi is not guaranteed to produce Delphic pronouncements. But overall, we believe law could benefit from more self-conscious use of the Delphi method—or at least the introduction of Delphi principles into analysis.

The most concerning potential drawbacks of Delphi are, in our view, that it (a) requires significant expertise in statistics, sampling, and empirical analysis, and (b) will not be helpful if done incorrectly.<sup>58</sup> But objection (b) is a criticism that can be made of any mode of analysis. Basic doctrinal analysis or even brute textual analysis (e.g., reading every word of a statute or rule literally) can easily be done poorly and will produce results at least as bad as even the most mishandled Delphi analysis.<sup>59</sup>

---

<sup>58</sup> See Gupta, *supra* note 41 (summarizing the different criticisms about the credibility of Delphi as a research method).

<sup>59</sup> One arguable difference is that poor doctrinal or textual analysis is limited to a single case, while the Delphi Method is more likely to be used for issues that span cases and are prospective in operation. However, this distinction is not so stark. When a single case is decided poorly as a result of bad doctrinal or textual analysis, it becomes a precedent that will be invoked for subsequent cases. In an adversary system, there will almost always be a self-interested party wishing to deploy even the most poorly reasoned case on its behalf. And because precedent is important, particularly in Anglo-American systems, the errors of a single doctrinal or textual decision will typically have ripple effects far beyond the instant matter.

Critics of Delphi might counter that precedent may not always be binding (e.g., a decision from another judicial system or a co-equal court rather than a higher court) and that it is easier for a court to overturn bad precedent than for a legislature to amend a problematic law or rule. The former observation, although valid, does not differentiate Delphi much from the force of precedent. That latter—which is ultimately an empirical question—remains unknown. Even widely reviled precedents often remain “good law” for decades while legislation may be revisited and improved frequently or rarely, depending on the political climate.

To work properly, the Delphi Method requires a skilled design and facilitator selecting a panel of true experts. Done incorrectly, Delphi could resemble a bad night of U.S. cable network news in which an “analysis” of a problem consists of hearing the opposing views of two (or more) partisans based on ideology and posturing rather than evidence.

But just as the average cable new broadcast could benefit from obtaining an injection of true expertise (e.g., having as guests genuine experts who have studied an issue for decades rather than two opposing political consultants, two opposing members of the legislature, or representatives of two opposing ideological think tanks), a well-executed Delphi study requires identification and recruitment of bona fide experts. To be sure, experts will have differing opinions, many of them formed by ideology and philosophy as well as education, training, and empirical research. But through good-faith exploration by the group, reaching common ground and effective problem-solving is a possibility.

Contrast the Delphi Method, which admittedly has some cumbersome features and is vulnerable to defective panel selection, response bias, or flawed questionnaires, with the norms of legal policymaking and adjudication. When the U.S. Congress, despite all its resources, considers legislation, it receives the bulk of its input from interest groups or partisan sources. Although academic and business experts are usually not completely absent, they take a back seat to lobbyists and electioneering. And too often, the legislators are not honest brokers attempting to find the correct solution but are acting in the service of an ideologically or electorally driven agenda set by their political parties, contributors, or select powerful constituents. Compared to this status quo, the Delphi Method could hardly do worse.

Expert panelists assessing this issue are far more likely to arrive at a consensus or near-consensus view regarding problems and possible solutions than lobbyists and partisans fighting to please portions of their electoral coalitions in a manner that provides the best sound-bites for re-election campaign advertisements.

Because legislation and administrative regulation have far-reaching impact, they justify extensive investment in analysis. In short, they are worth the time and cost of assessing via the Delphi Method or its substantial equivalent. Adjudication presents a more difficult cost-benefit problem.

At one end of the spectrum are isolated cases that are largely limited to particular facts. In these situations, attempts to use the Delphi Method, if apt at all, would not pass the cost-benefit test. At the other end of the litigation spectrum, however, are very large cases (e.g., a government antitrust decision to break up an alleged monopoly or prohibit or permit a merger) that would almost certainly be Delphi-worthy.

For example, if the issue is whether permitting two large pharmaceutical companies to merge will result or in injury or advantage to consumers or an increase or decline in research, innovation, and development, a panel of experts participating in a Delphi inquiry is almost certainly better than the current spectacle

of duelling economists hired by opposing sides being assessed by a judge or lay jurors lacking any formal economic training or business experience.

In similar fashion, a case of first impression in which a court's decision will as a practical matter have extensive precedential influence is also a strong candidate for Delphi inquiry. Should persons suffering emotional distress be considered to have suffered "bodily injury" within the meaning of a liability insurance policy? The question is not only philosophical, but also concerns whether treating such a claim as within the policy will result in a large increase in litigation, more fabricated injuries, higher insurance premiums and other unintended consequences. Also of concern is whether failing to cover such claims will leave liability insurance policyholders unduly injured by uninsured liability, whether as a result taxpayer-funded social programs will fill the gap, or whether victims of torts are undercompensated.

Courts have done an adequate, but not perfect, job of addressing these concerns—but they have largely done so based on judges' own personal views of likely outcomes without benefit of any real study or expertise on the topic.

As one travels on the continuum from a "big case" or "precedent-setting case" toward an individual or low-stakes dispute, the case for investing in the Delphi Method wanes, but it does not dissipate. The Delphi Method is particularly suited to assessing issues that can be addressed by statute or administrative regulation. But it can be a valuable addition to adjudication, as well. And in some instance, careful work has been done by experts in a manner resembling Delphi.

For example, the September 2016 PCAST Report to the President of the United States,<sup>60</sup> although not intentionally following the precise Delphi methodology, formed a task force of experts and support staff that carefully canvassed literature of the forensic sciences.<sup>61</sup> The result was an assessment by neutral experts examining, and to a large degree debunking several popular evidentiary devices: DNA matching, fingerprinting, ballistics, and bite mark evidence.<sup>62</sup> Noting the limitations of this evidence (and in the case of bite marks, its rather stark unreliability), the PCAST Report then offered a number of guidelines to assist courts in handling such evidence, providing a range of options ranging from exclusion to limited admissibility to cautionary instructions to simply being careful not to imbue the evidence with (we wince, but have to say it) "Delphic" authority, which can make jurors (and even judges) too starstruck by pseudo-science or information that, although legitimately evidentiary, should not be viewed as conclusive.<sup>63</sup>

---

<sup>60</sup> See PCAST REPORT, *supra* note 45.

<sup>61</sup> See *id.* at iv-xi (noting membership of Council, topic of the Report, and describing methodology and investigation). PCAST describes itself as "an advisory group of the Nation's leading scientists and engineers, appointed by the President to augment the science and technology advice available to him from inside the White House and from cabinet departments and other Federal agencies."

<sup>62</sup> See *id.* at 69-83 (DNA analysis); see also, *id.* at 83-87 (bitemark analysis); *id.* at 87-103 (latent fingerprint analysis); *id.* at 104-14 (ballistics analysis); *id.* at 114-18 (footwear analysis); *id.* at 188-123 (hair analysis).

<sup>63</sup> See *id.* at 142-146. The PCAST Report also made recommendations to scientists, the FBI and the Attorney General in addressing forensic evidence going forward. See *id.* at 124-41.



We do not advocate over-use or inefficient use of the Delphi Method. Rather, we urge greater consideration of a Delphi inquiry in policymaking and adjudication and use of the Delphi Method in apt circumstances. It is a vehicle for analysis insufficiently appreciated and explored by the legal system. We do not recommend a Copernican or paradigm-shifting twist for the legal doctrine. The Delphi Method allows studies to be conducted which adopt the internal perspective of studying Law as its own discipline and offering results that jurists can well accept, but with more rigour as it provides a much more verified judgment of lawfulness.

The ambition of introducing a new methodology will no doubt clash with jurists' natural tendency to conservatism (in methodology, regardless of the individual judges' jurisprudential or political ideology), and even more so when we analyze the mathematical part it implies. It is easier to insult what is unknown than making the effort to be open to new methodological methods, despite them offering more scientific rigor than those that have so far been valid.

The Delphi Method is not a panacea. But, if properly deployed in apt situations, it holds substantial promise for improving the quality of analysis and decisionmaking in adjudication, administration, legislation, and the overall policy-making.

