

2017

## Inside the Arbitrator's Mind

Susan D. Franck

Follow this and additional works at: <https://scholarlycommons.law.emory.edu/elj>

---

### Recommended Citation

Susan D. Franck, *Inside the Arbitrator's Mind*, 66 Emory L. J. 1115 (2017).

Available at: <https://scholarlycommons.law.emory.edu/elj/vol66/iss5/2>

This Article is brought to you for free and open access by the Journals at Emory Law Scholarly Commons. It has been accepted for inclusion in Emory Law Journal by an authorized editor of Emory Law Scholarly Commons. For more information, please contact [law-scholarly-commons@emory.edu](mailto:law-scholarly-commons@emory.edu).

## INSIDE THE ARBITRATOR'S MIND

*Susan D. Franck*

*Anne van Aaken*

*James Freda*

*Chris Guthrie*

*Jeffrey J. Rachlinski\**

*Arbitrators are lead actors in global dispute resolution. They are to global dispute resolution what judges are to domestic dispute resolution. Despite this, arbitral decisionmaking is a black box. This Article is the first to use original experimental research to explore how international arbitrators decide cases. We find that arbitrators often make intuitive and impressionistic decisions rather than fully deliberative ones. We also find evidence that casts doubt on the conventional wisdom that arbitrators render “split the baby” decisions. Although direct comparisons are difficult, we find that arbitrators generally perform at least as well as, but never demonstrably worse than, national judges analyzed in earlier research. There may be reasons to prefer judges to international arbitrators, but the quality of judgment and decisionmaking, at least as measured in these experimental studies, is not one of them. Thus, normative debates about global dispute resolution should focus not on*

---

\* Susan D. Franck is a Professor of Law, at American University, Washington College of Law. Anne van Aaken is the Professor of Law and Economics, Legal Theory, Public International Law and European Law, University of St. Gallen. James Freda is an attorney and diplomat at the United Nations; the views expressed in this Article are solely those of the authors and do not reflect the views of the United Nations. Chris Guthrie is the Dean and John Wade-Kent Syverud Professor of Law at Vanderbilt Law School. Jeffrey Rachlinski is the Henry Allen Mark Professor of Law at Cornell University Law School. This scholarship benefited from presentations at American University's Washington College of Law, Bar Ilan University, Columbia Law School, Fordham Law School, Seton Hall Law School, St. John's Law School, Texas A&M Law School, Washington & Lee University School of Law, the European Society of International Law, the University of London Queen Mary Conference on Arbitration and Legal Reasoning, and comments by Robert Ahdieh, José Alvarez, George Bermann, Chris Brummer, Miriam Cherry, Julian Davis Mortenson, William Dodge, Howard Erichson, Christopher Drahozal, Jean Galbraith, Alexandra Klein, Irina Manta, Jacqueline Nolan-Haley, W. Michael Reisman, Jonathan Romberg, and Brian Sheppard. We are grateful to Lucy Reed who was bold enough to support our unusual research. We thank the ICCA Miami Congress host committee who provided logistical support and all of the participants who took the research seriously and generously gave their time; and we thank John Barkett and Shook, Hardy & Bacon LLP who permitted us to use their offices for initial coding during the conference. The diligent coding of Trista Bishop-Watt, Stephen Halpin, Sharon Jeong, Rachael Kurzweil, Kellen Lavin, Tobias Lehmann, George Mackie, Bret Marfut, Stephanie Miller, Krystal Swendsboe, and the support of Washington & Lee Law Librarians, Caroline Osborne and Stephanie Miller, made our research possible. The Washington & Lee Frances Lewis Law Center, Washington & Lee Transnational Law Institute, and University of St. Gallen Law School provided research support.

*decisionmaker identity or title but rather on structural safeguards and legal protections to enhance quality rule of law based decisionmaking.*

## INTRODUCTION

Arbitration is an important alternative to litigation in the United States, particularly in consumer, employment, and securities disputes.<sup>1</sup> But arbitration's role in domestic dispute resolution pales in comparison to the role it plays globally. In most international disputes, arbitration is the default dispute resolution method.<sup>2</sup>

This means that arbitrators are the central actors in international dispute resolution. They play a vital role in the global economy, oversee disputes involving billions of dollars, and make decisions implicating the transnational rule of law.

Despite the outsized role that arbitrators play in international dispute resolution, we know relatively little about how they make decisions. Some commentators sing arbitrators' praises,<sup>3</sup> observing that they possess both subject-matter expertise and incentives to resolve disputes according to governing law. Other commentators decry their skill and demand instead that judges resolve disputes.<sup>4</sup> They question the quality of arbitrator

---

<sup>1</sup> See *infra* note 22 and accompanying text.

<sup>2</sup> George A. Bermann, *International Commercial Arbitration: Past, Present, Future*, 33 ALTERNATIVES TO HIGH COST LITIG. (Int'l Inst. for Conflict Prevention & Resolution), May 2015, at 65, 65; see also Gilles Cuniberti, *Beyond Contract—The Case for Default Arbitration in International Commercial Disputes*, 32 FORDHAM INT'L L.J. 417, 417–18 (2009); Christopher R. Drahozal, *New Experiences of International Arbitration in the United States*, 54 AM. J. COMP. L. 233, 233 (2006) (“Between 1993 and 2003, the number of international arbitration proceedings administered by leading institutions almost doubled.”); Stephen R. Halpin III, *Stayin’ Alive?: BG Group, PLC v. Republic of Argentina and the Vitality of Host-Country Litigation Requirements in Investment Treaty Arbitration*, 71 WASH. & LEE L. REV. 1979, 2021–22 (2014) (“[I]nternational arbitration between foreign investors and host countries will remain the dominant method of conclusively resolving investment disputes . . .”).

<sup>3</sup> See, e.g., Andreas F. Lowenfeld, *The Elements of Procedure: Are They Separately Portable?*, 45 AM. J. COMP. L. 649, 654 (1997); see also Catherine A. Rogers, *The Vocation of the International Arbitrator*, 20 AM. U. INT'L L. REV. 957, 958–59 (2005); Jason Webb Yackee, *Investment Treaties and Investor Corruption: An Emerging Defense for Host States*, 52 VA. J. INT'L L. 723, 744 n.105 (2012).

<sup>4</sup> Letter from Alliance for Justice to U.S. Congressional Officials and U.S. Trade Representative (Mar. 11, 2015), <http://www.afj.org/wp-content/uploads/2015/03/ISDS-Letter-3.11.pdf>; see also *supra* note 76 and accompanying text (identifying that German judges publicly rejected arbitration of international investment disputes and demanded disputes be returned to national courts). Other critiques of international arbitration address transparency, review by national courts, consistency in outcome, and diversity of adjudicators. Susan D. Franck et al., *The Diversity Challenge: Exploring the “Invisible College” of International Arbitration*, 53 COLUM. J. TRANSNAT'L L. 429 (2015); Susan D. Franck, *The Legitimacy Crisis in Investment Treaty Arbitration:*

decisionmaking,<sup>5</sup> arguing that arbitrators often ignore applicable law<sup>6</sup> and generally “split the baby” by making awards that fall halfway between the positions the parties advance.<sup>7</sup> Whatever perspective they espouse, commentators debate the relative merits of international arbitration in an information vacuum.

In an effort to shed some light on arbitration, this Article reports the results of a first-ever set of experiments involving international arbitrator decisionmaking.<sup>8</sup> In it, we describe how international arbitrators decide hypothetical cases. When possible, we compare arbitrators' performance to domestic judges. We also explore how the experimental insights we glean might inform adjudicative design.

To do so, we draw on decades of experimental research on the psychology of judgment and decisionmaking. That research shows—contrary to the assumptions of classical economics but consistent with common sense—that

---

*Privatizing Public International Law Through Inconsistent Decisions*, 73 *FORDHAM L. REV.* 1521 (2005). These concerns are beyond this Article, as they do not address decision-making psychology.

<sup>5</sup> See Thomas J. Stipanowich, *Reflections on the State and Future of Commercial Arbitration: Challenges, Opportunities, Proposals*, 25 *AM. REV. INT'L ARB.* 297, 361 (2014); see also Tom Ginsburg, *The Arbitrator as Agent: Why Deferential Review Is Not Always Pro-Arbitration*, 77 *U. CHI. L. REV.* 1013, 1014 (2010) (“[A]rbitrators might deliver poor-quality decisions that undermine the attractiveness of arbitration as a whole.”); Thomas J. Stipanowich, *Rethinking American Arbitration*, 63 *IND. L.J.* 425, 458 (1988) (“[T]he less favorable a person's view of the quality of decisionmakers in arbitration, the more likely that person was to support broader judicial review of arbitration awards.”).

<sup>6</sup> See David S. Baffa, John L. Collins & Gerald L. Maatman, Jr., *Guidance for Employers Considering Mandatory Arbitration Agreements with Class and Collective Action Waivers*, 39 *EMP. REL. L.J.* 34, 41 (2013); see also Henry Wade Rogers, *The Essentials of a Law Establishing an International Court*, 22 *YALE L.J.* 277, 287 (1913) (“[O]ne who carefully examines the decisions rendered by the Arbitral Tribunals will come to the conclusion that they are inferior to those rendered in the Supreme Court of the United States.”); Peter B. Rutledge, *Toward a Contractual Approach for Arbitral Immunity*, 39 *GA. L. REV.* 151, 175 (2004) (observing that immunity “allows arbitrators to render poor or unenforceable decisions and then . . . escape responsibility”).

<sup>7</sup> William W. Park, *Arbitrator Integrity: The Transient and the Permanent*, 46 *SAN DIEGO L. REV.* 629, 689–93 (2009); Anthea Roberts, *Clash of Paradigms: Actors and Analogies Shaping the Investment Treaty System*, 107 *AM. J. INT'L L.* 45, 93 (2013); see also WILLIAM W. PARK, *ARBITRATION OF INTERNATIONAL BUSINESS DISPUTES: STUDIES IN LAW AND PRACTICE* 560 (2d ed. 2012) (describing bankers' herd mentality and suggesting arbitration in an unnecessary invitation to render split the difference awards); Richard A. Posner, *Judicial Behavior and Performance: An Economic Approach*, 32 *FL. ST. L. REV.* 1259, 1261 (2005) (“We can expect, therefore, a tendency for arbitrators to ‘split the difference’ in their awards . . . .”); Joshua B. Simmons, *Valuation in Investor-State Arbitration: Toward a More Exact Science*, 30 *BERKELEY J. INT'L L.* 196, 200, 208–14 (2012) (identifying “perceptions that arbitrators merely ‘split the baby’ between the parties’ proposed valuations, particularly when awards are poorly explained”).

<sup>8</sup> *But see infra* note 83 and accompanying text (describing how, until recently, most exploration about cognitive illusions in international arbitration was largely theoretical).

human beings often make decisions in irrational, but predictable, ways.<sup>9</sup> Likewise, we draw on more recent research showing that judges, like other human beings, are also prone to predictably irrational decisionmaking.<sup>10</sup> But what about arbitrators?

We might hypothesize that arbitrators make decisions much like judges. Arbitrators, like judges, are human beings; both arbitrators and judges are elite professionals engaged in the task of applying legal principals to facts and have a legal mandate to exercise their judgment in a neutral and objective way. On the other hand, we might hypothesize that arbitrators and judges make decisions differently, as each have different incentives, mandates from different principals,<sup>11</sup> different cultures and legal traditions,<sup>12</sup> and different subject matter

---

<sup>9</sup> See generally DAN ARIELY, PREDICTABLY IRRATIONAL: THE HIDDEN FORCES THAT SHAPE OUR DECISIONS (rev. & expanded 2008); DANIEL KAHNEMAN, THINKING, FAST AND SLOW (2011); SCOTT PLOUS, THE PSYCHOLOGY OF JUDGMENT AND DECISION MAKING (1993).

<sup>10</sup> Initial research on cognition and judicial decisionmaking used the term “cognitive illusions” to describe intuitive, simple, quick assessments. Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 782 (2001); see also *infra* note 77. Psychologists and behavioral economists call these “biases and heuristics.” In international arbitration, “bias” has a loaded, often undefined, meaning, whereas “independence” and “impartiality” have precise legal meanings. See MARGARET L. MOSES, THE PRINCIPLES AND PRACTICE OF INTERNATIONAL COMMERCIAL ARBITRATION 135–36 (2d ed. 2012); Dominique Hascher, *Independence and Impartiality of Arbitrators: 3 Issues*, 27 AM. U. INT’L L. REV. 789, 791–92 (2012); *infra* note 70 and accompanying text. We use “cognitive illusion” to avoid confusion and to focus on intuitive cognition.

<sup>11</sup> States sometimes appoint judges to long-term appointments with a broad mandate; other times, national judges are elected or have finite jurisdiction. See, e.g., APPOINTING JUDGES IN AN AGE OF JUDICIAL POWER: CRITICAL PERSPECTIVES FROM AROUND THE WORLD (Kate Malleon & Peter H. Russell eds., 2006). By contrast, parties appoint arbitrators, although courts or institutions can also appoint arbitrators. GARY B. BORN, INTERNATIONAL COMMERCIAL ARBITRATION: COMMENTARY AND MATERIALS 629 (2d ed. 2001); *infra* notes 69–70. States pay judges; but parties pay arbitrators, and tribunals allocate costs. Susan D. Franck, *Rationalizing Costs in Investment Treaty Arbitration*, 88 WASH. U. L. REV. 769 (2011); see also Ethan J. Leib, David L. Ponet & Michael Serota, *A Fiduciary Theory of Judging*, 101 CALIF. L. REV. 699, 722 (2013) (arguing judges have a fiduciary duty to the legislature and public in some cases but “arbitrators do not hold the judicial office in a democracy and therefore do not have a responsibility to the people in the way judges do”). Arbitrators may have financial interests in re-appointment given prospects of further income, but arbitrators have other incentives like reputation or lost opportunity of pursuing work that is more fiscally lucrative or less likely to create conflicts of interest. See Robert O. Keohane, *Rational Choice Theory and International Law: Insights and Limitations*, 31 J. LEGAL STUD. 307, 309 (2002) (“[I]t is important not to equate rationality with materialistic self-interest . . .”).

<sup>12</sup> See KAREN J. ALTER, THE NEW TERRAIN OF INTERNATIONAL LAW: COURTS, POLITICS, RIGHTS, at xix–xx (2014) (observing domestic adjudicators may have different approaches than international courts and tribunals); Posner, *supra* note 7, at 1259 (“[J]udicial behavior is likely to differ across national legal systems and indeed within a nation’s legal systems . . .”); Leon E. Trakman, “Legal Traditions” and International Commercial Arbitration, 17 AM. REV. INT’L ARB. 1, 2–3 (2006) (discussing different cultures within international arbitration); Vitalius Tumonis, *Adjudication Fallacies: The Role of International Courts in Interstate Dispute Settlement*, 31 WISC. INT’L L.J. 35, 36 (2013) (noting the “fallacy” that “international courts are essentially analogous to their domestic counterparts, when in fact there are many more differences between them than similarities”).

expertise. Judges, as generalists, may be relatively unfamiliar with the facts, law, and context of a case in front of them; arbitrators, by contrast, often have highly relevant domain expertise.<sup>13</sup>

Understanding how arbitrators decide is important because it can inform hotly contested debates over the proper forms of dispute resolution to deploy both international and national disputes. Senator Elizabeth Warren has taken issue with the use of arbitration and objected to the lack of “independent judges”<sup>14</sup> in the Trans-Pacific Partnership.<sup>15</sup> Likewise, the European Parliament expressed a desire to strip arbitrators of jurisdiction in trade agreements<sup>16</sup> with the United States, namely the Trans-Atlantic Trade and Investment Partnership (TTIP), and with Canada, namely, the Comprehensive Economic and Trade Agreement (CETA);<sup>17</sup> instead, the EU demands that judges must resolve

---

<sup>13</sup> See, e.g., Christopher R. Drahozal, *Private Ordering and International Commercial Arbitration*, 113 PENN ST. L. REV 1031, 1046 (2009); see also *BG Group PLC v. Republic of Argentina*, 134 S. Ct. 1198, 1210 (2014) (“International arbitrators are likely more familiar than are judges with the expectations of foreign investors and recipient nations . . . .”); *Mitsubishi Motors Corp. v. Soler Chrysler-Plymouth, Inc.*, 473 U.S. 614, 633–34 (1985) (noting the specialist, elite international arbitrators appointed in that case).

<sup>14</sup> *TPP Opponents, Warren, Academics Highlight ISDS As Key Reason to Resist Deal*, INSIDE U.S. TRADE (Sept. 8, 2016), <https://insidetrade.com/daily-news/tpp-opponents-warren-academics-highlight-isds-key-reason-resist-deal>; Elizabeth Warren, *The Trans-Pacific Partnership Clause Everyone Should Oppose*, WASH. POST (Feb. 25, 2015), [https://www.washingtonpost.com/opinions/kill-the-dispute-settlement-language-in-the-trans-pacific-partnership/2015/02/25/ec7705a2-bd1e-11e4-b274-e5209a3bc9a9\\_story.html](https://www.washingtonpost.com/opinions/kill-the-dispute-settlement-language-in-the-trans-pacific-partnership/2015/02/25/ec7705a2-bd1e-11e4-b274-e5209a3bc9a9_story.html).

<sup>15</sup> At the time of writing this article, TPP was signed and was moving forward towards enactment into law but its future was somewhat uncertain. Compare Tim Worstall, *With Trump's Election the TPP Probably Is Dead, Yes—As Is the TTIP*, FORBES (Nov. 11, 2016, 4:35 AM), <http://www.forbes.com/sites/timworstall/2016/11/11/with-trumps-election-the-tpp-probably-is-dead-yes-as-is-the-ttip/#5104e7185b80> (postulating after President Trump's election that the TPP would not survive), and Mike DeBonis, Ed O'Keefe & Ana Swanson, *The Trans-Pacific Partnership Is Dead, Schumer Tells Labor Leaders*, WASH. POST (Nov. 10, 2016), [https://www.washingtonpost.com/news/powerpost/wp/2016/11/10/the-trans-pacific-partnership-is-dead-schumer-tells-labor-leaders/?utm\\_term=.9fc6c62d1d98](https://www.washingtonpost.com/news/powerpost/wp/2016/11/10/the-trans-pacific-partnership-is-dead-schumer-tells-labor-leaders/?utm_term=.9fc6c62d1d98) (discussing senators' beliefs that the TPP will not pass in Congress), with Cyrus Sanati, *Trans-Pacific Partnership May Not Be Dead Yet*, USA TODAY (Nov. 21, 2016, 8:07 AM), <http://www.usatoday.com/story/tech/columnist/2016/11/20/trans-pacific-partnership-may-not-dead-yet/93986892/> (discussing the benefits of the TPP and expressing belief that it may survive). During the midst of editing, an Executive Order withdrew the United States from the TPP. *Trump Signs EO Removing US from TPP*, C-SPAN (Jan. 23, 2017), <https://www.c-span.org/video/?c4651802/trump-ao-tpp&start=24>.

<sup>16</sup> See *EU Finalizes Proposal for Investment Protection and Court System for TTIP*, EUROPEAN COMM'N (Nov. 12, 2015), <http://trade.ec.europa.eu/doclib/press/index.cfm?id=1396>; see also EU TTIP Team (@EU\_TTIP\_Team), TWITTER (Sept. 16, 2015, 4:30 AM), [https://twitter.com/EU\\_TTIP\\_team/status/644110990242639873](https://twitter.com/EU_TTIP_team/status/644110990242639873).

<sup>17</sup> The original, signed version of CETA included arbitration; but in an unprecedented “scrubbing” process, arbitration was replaced wholesale with a standing court. Wolfgang Alschner, *Legal Scrubbing or Renegotiation? A Text-as-Data Analysis of How the EU Smuggled an Investment Court into Its Trade Agreement with Canada*, MAPPING BITS BLOG (Mar. 24, 2016), <http://mappinginvestmenttreaties.com/blog/2016/03/legal%20scrubbing-ceta/>. While drafting this Article, there were ongoing concerns as to whether CETA will have any force and effect. Kathleen Harris, *Justin Trudeau Says CETA Will Test European Union's 'Usefulness'*,

disputes.<sup>18</sup> There are similar concerns about arbitrators' suitability to decide wholly domestic disputes.<sup>19</sup>

This Article—in which we peer inside the arbitral mind—aspires to offer an objective, empirical, and evidence-based approach to these important normative choices about transnational dispute system design. Ultimate design choices are part of a larger puzzle that will inevitably be affected by multiple variables,<sup>20</sup> including practical politics, political economy, and norm preferences.<sup>21</sup> But by focusing on arbitrator cognition and competence, we hope to contribute to these design choices.

In Part I of the Article, we introduce international arbitration and behavioral psychology. In Part II, we identify our hypotheses and experimental

CBC NEWS (Oct. 13, 2016, 2:55 PM), <http://www.cbc.ca/news/politics/manuel-valls-parliament-hill-trudeau-1.3802584>. Likewise, with the Brexit vote, TTIP negotiations are stalled. Jim Zarroli, *German Official Says U.S.-Europe Trade Talks Have Collapsed, Blames Washington*, NPR (Aug. 26, 2016, 4:31 PM), <http://www.npr.org/sections/thetwo-way/2016/08/28/491721332/german-official-says-u-s-europe-trade-talks-have-collapsed-blames-washington>.

<sup>18</sup> Recently, the EU appears to have moved towards creating a multilateral, rather than a series of bilateral, investment courts. See, e.g., *Inception Impact Assessment*, EUROPEAN COMM'N (Aug. 1, 2016), [http://ec.europa.eu/smart-regulation/roadmaps/docs/2016\\_trade\\_024\\_court\\_on\\_investment\\_en.pdf](http://ec.europa.eu/smart-regulation/roadmaps/docs/2016_trade_024_court_on_investment_en.pdf) (outlining the process for moving forward with a multilateral investment court); EUROPEAN COMM'N, CONSULTATION STRATEGY, IMPACT ASSESSMENT ON THE ESTABLISHMENT OF A MULTILATERAL INVESTMENT COURT FOR INVESTMENT DISPUTE RESOLUTION, (2016), [http://trade.ec.europa.eu/doclib/docs/2016/october/tradoc\\_154997.09.30%20Consultation%20strategy%20IIA\\_for%20publication.pdf](http://trade.ec.europa.eu/doclib/docs/2016/october/tradoc_154997.09.30%20Consultation%20strategy%20IIA_for%20publication.pdf) (outlining the process for moving forward with a multilateral investment court).

<sup>19</sup> See, e.g., Jessica Silver-Greenberg & Michael Corkery, *In Arbitration, a 'Privatization of the Justice System'*, N.Y. TIMES (Nov. 1, 2015) <https://www.nytimes.com/2015/11/02/business/dealbook/in-arbitration-a-privatization-of-the-justice-system.html?ref=topics>; Jessica Silver-Greenberg & Robert Gebeloff, *Arbitration Everywhere, Stacking the Deck of Justice*, N.Y. TIMES (Oct. 31, 2015), [https://www.nytimes.com/2015/11/01/business/dealbook/arbitration-everywhere-stacking-the-deck-of-justice.html?\\_r=1](https://www.nytimes.com/2015/11/01/business/dealbook/arbitration-everywhere-stacking-the-deck-of-justice.html?_r=1); The Editorial Board, *Arbitrating Disputes, Denying Justice*, N.Y. TIMES (Nov. 7, 2015), <https://www.nytimes.com/2015/11/08/opinion/sunday/arbitrating-disputes-denying-justice.html?ref=topics>.

<sup>20</sup> See generally Amy J. Cohen, *Dispute Systems Design, Neoliberalism, and the Problem of Scale*, 14 HARV. NEGOT. L. REV. 51 (2009); Susan D. Franck, *Integrating Investment Treaty Conflict and Dispute Systems Design*, 92 MINN. L. REV. 161 (2007). While aspects of this Article compare arbitration and litigation, other processes—including negotiation and mediation—are core options in system design and promote norms like distributive and procedural justice. Carrie Menkel-Meadow, *Are There Systemic Ethics Issues in Dispute System Design? And What We Should [Not] Do About It: Lessons from International and Domestic Fronts*, 14 HARV. NEGOT. L. REV. 195 (2009); Andrea Kupfer Schneider, *The Intersection of Dispute Systems Design and Transitional Justice*, 14 HARV. NEGOT. L. REV. 289 (2009).

<sup>21</sup> Other factors, beyond those in our experiment, invariably influence system design. These might include concerns related to certainty, predictability, transparency, conflicts of interest, impartiality, legal correctness, efficiency, enforceability, and diversity. See *supra* note 4 and note 11 (identifying arbitration-related concerns); see also *infra* note 70 and note 244 (identifying arbitration-related concerns). We do not address conflicts of interest or impartiality in real disputes, as those subjects are better analyzed through arbitration doctrine or content analysis of existing cases.

methodology. In Part III, we report experimental results showing that arbitrators, like judges, are prone to intuitive decisionmaking and the influence of well-known cognitive illusions like anchoring, framing, representativeness, and egocentrism. In Part IV, we interpret the results, acknowledge the limitations of our study, and offer normative assessments. Recognizing that intuition influences adjudicative determinations irrespective of an adjudicator's title or mandate, we argue that dispute system designers should not focus on whether judges or arbitrators should decide disputes. Rather, system designers should focus on structural and procedural reforms to decrease the risk of error and to promote quality decisionmaking in international economic dispute settlement.

## I. INTERNATIONAL ARBITRATION

Arbitration is a ubiquitous method of dispute settlement used in both domestic<sup>22</sup> and international disputes.<sup>23</sup> This section explores the prevalence and vitality of international arbitration. It introduces international commercial arbitration (ICA) and international treaty arbitration (ITA), explains arbitral procedures, and discusses arbitral decisionmaking.

### A. *A Doctrinal Primer*

Parties involved in global economic activity require reliable dispute resolution. Although parties can use informal processes—like negotiation or mediation—these methods operate in the “shadow of the law.”<sup>24</sup> International

---

<sup>22</sup> U.S. domestic arbitration involves consumer, employment, franchise, and securities law disputes. Stephen J. Ware, *Teaching Arbitration Law*, 14 AM. REV. INT'L ARB. 231, 239 (2003); see also CONSUMER FIN. PROT. BUREAU, ARBITRATION STUDY, § 5, 30 (2015); Sarah Rudolph Cole, *The Federalization of Consumer Arbitration: Possible Solutions*, 2013 U. CHI. LEGAL F. 271, 272–75; Christopher R. Drahozal & Quentin R. Wittrock, *Is There a Flight from Arbitration?*, 37 HOFSTRA L. REV. 71, 74–75 (2008); Jill I. Gross, *The End of Mandatory Securities Arbitration?*, 30 PACE L. REV. 1174, 1176–77 (2010); Constantine Katsoris, *Securities Arbitrators Do Not Grow on Trees*, 14 FORDHAM J. CORP. & FIN. L. 49, 50 (2008); Erin O'Hara O'Connor, Kenneth J. Martin & Randall S. Thomas, *Customizing Employment Arbitration*, 98 IOWA L. REV. 133 (2012). Employment arbitration is distinguishable from labor arbitration with a distinct doctrinal regime. Arthur T. Carter, Edward F. Berbarie & Sean M. McCrory, *The Principal Differences Between Labor and Employment Arbitration*, 69 THE ADVOCATE 85 (2014); William B. Gould IV, *Kissing Cousins?: The Federal Arbitration Act and Modern Labor Arbitration*, 55 EMORY L.J. 609 (2006).

<sup>23</sup> Internationally, arbitration offers a proxy for diplomatic negotiation or state-to-state dispute settlement. Susan D. Franck, *Foreword: A Symposium Exploring the Modern Legacy of William Jennings Bryan*, 86 NEB. L. REV. 142, 144–45 (2007); Anthea Roberts, *State-to-State Investment Treaty Arbitration: A Hybrid Theory of Interdependent Rights and Shared Interpretive Authority*, 55 HARV. INT'L L.J. 1 (2014).

<sup>24</sup> Robert Cooter, Stephen Marks & Robert Mnookin, *Bargaining in the Shadow of the Law: A Testable Model of Strategic Behavior*, 11 J. LEGAL STUD. 225 (1982). See generally Robert H. Mnookin & Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case of Divorce*, 88 YALE L.J. 950 (1979).



arbitration offers formal adjudication to provide a final and binding assessment of legal rights and “has achieved a level of legitimacy to which other [types of international] disciplines can only aspire.”<sup>25</sup>

There are multiple reasons international arbitration enjoys this stature, including historic pedigree,<sup>26</sup> adaptability to new contexts and the flexibility of the process,<sup>27</sup> the capacity to provide neutrality while avoiding fears that locals will be favored over foreigners,<sup>28</sup> expertise, and a strong enforcement record.<sup>29</sup> Moreover, in low-capacity environments where court systems may be weak, international arbitration fills a crucial developmental gap.<sup>30</sup>

Two core areas of modern international arbitration are ICA and ITA.<sup>31</sup> ICA is a traditional form of arbitration where parties resolve transnational disputes under national law.<sup>32</sup> ICA covers a broad range of disputes, including contract breach, business torts, and antitrust violations.<sup>33</sup> It typically involves commercial disputes between two businesses, but it can also encompass contract disputes between investors and states under national law related to commercial

---

<sup>25</sup> S.I. Strong, *Beyond the Self-Execution Analysis: Rationalizing Constitutional, Treaty, and Statutory Interpretation in International Commercial Arbitration*, 53 VA. J. INT'L L. 499, 572–73 (2013).

<sup>26</sup> Sabra A. Jones, *Historical Development of Commercial Arbitration in the United States*, 12 MINN. L. REV. 240, 242–43 (1928); Earl S. Wolaver, *The Historical Background of Commercial Arbitration*, 83 U. PA. L. REV. 132, 132 (1934); see also *infra* notes 40–42.

<sup>27</sup> See *infra* notes 31, 62–70 (describing international arbitration’s substantive and procedural flexibility).

<sup>28</sup> See George A. Bermann et al., *Restating the U.S. Law of International Commercial Arbitration*, 113 PENN. ST. L. REV. 1333, 1342 (2009) (“Parties choose international arbitration primarily because they fear being subject to the potentially biased decisions of the national courts of their business-partner-turned-adversary.”); Loukas Mistelis & Crina Baltag, *Recognition and Enforcement of Arbitral Awards and Settlement in International Arbitration: Corporate Attitudes and Practices*, 19 AM. REV. INT’L ARB. 319, 320 (2008) (“[G]rowth of arbitration has been driven by flaws in the national legal systems and the distrust and suspicion associated with litigation in a foreign country . . .”).

<sup>29</sup> HERBERT KRONKE, INTRODUCTION TO RECOGNITION AND ENFORCEMENT OF FOREIGN ARBITRAL AWARDS: A GLOBAL COMMENTARY ON THE NEW YORK CONVENTION 1, 3 (Herbert Kronke et al. eds., 2010); see also *infra* note 73 (describing enforcement).

<sup>30</sup> Mistelis & Baltag, *supra* note 28, at 320–21.

<sup>31</sup> Ban-Ki Moon, UN Secretary-General Ban-ki Moon’s Address to ICCA 2016 Congress, May 9, 2016, <https://www.un.org/sg/en/content/sg/statement/2016-05-09/secretary-generals-address-international-council-commercial> («l’arbitrage peut jouer un rôle clef pour ce qui est de restaurer l’état de droit après un conflit, puisqu’établir un système judiciaire pleinement indépendant peut prendre du temps»). Arbitration extends to other areas. See, e.g., *supra* note 22.

<sup>32</sup> Depending upon the applicable law, ICA may require application of transnational, including the Convention on the International Sale of Goods (CISG) UNIDROIT, law principles. George Bermann, *Restating the U.S. Law of International Commercial Arbitration*, 42 N.Y.U. J. INT’L L. & POL. 175, 190–91 (2009).

<sup>33</sup> See, e.g., JULIAN D. M. LEW, LOUKAS MISTELIS & STEFAN M. KRÖLL, COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION 187–219 (2003); JEAN-FRANÇOIS POUURET & SÉBASTIEN BESSON, COMPARATIVE LAW OF INTERNATIONAL ARBITRATION 265–73 (2007).

ventures or infrastructure projects.<sup>34</sup> Arbitrators use existing commercial law—whether codified in national statutes, case law, or otherwise—to adjudicate claims and finally resolve disputes.<sup>35</sup>

ICA is common and growing.<sup>36</sup> International arbitration centers report that hundreds of cases are filed annually.<sup>37</sup> Commentators identified over 2700 disputes involved in institutional arbitration in one year and “major claims” involving billions of dollars.<sup>38</sup> In its 2015 analysis, the *American Lawyer* identified over 125 cases with billion-dollar claims.<sup>39</sup>

ITA, or arbitration devolving from international law-based treaty rights states grant to investors, also has deep roots. It arose from international law mixed-claims commissions where states created sui generis opportunities for private individuals or entities to bring claims against states for economic harm.

---

<sup>34</sup> HEGE ELISABETH KJOS, *APPLICABLE LAW IN INVESTOR-STATE ARBITRATION* 158, 163 (Vaughan Lowe, Dan Sarooshi & Stefan Talmon eds., 2013); Lise Johnson & Oleksandr Volkov, *Investor-State Contracts, Host-State “Commitments” and the Myth of Stability in International Law*, 24 *AM. REV. INT’L ARB.* 361, 382–83 (2013).

<sup>35</sup> W. Laurence Craig, *The Arbitrator’s Mission and the Application of Law in International Commercial Arbitration*, 21 *AM. REV. INT’L ARB.* 243, 260 (2010); see Susan D. Franck, *The Role of International Arbitrators*, 12 *ILSA J. INT’L & COMP. L.* 499, 504 (2006). It is possible to apply more nebulous conceptions of fairness, namely principles of *amiable compositeur* or *ex aequo et bono*; but this is uncommon and requires parties to opt-in to the discretion. *Id.*; Leon Trakman, *Ex Aequo Et Bono: Demystifying an Ancient Concept*, 8 *CHI. J. INT’L L.* 621, 623, 632 n.64 (2008).

<sup>36</sup> TOWARDS A SCIENCE OF INTERNATIONAL ARBITRATION: COLLECTED EMPIRICAL RESEARCH 341 app. 1 (Christopher R. Drahozal & Richard W. Naimark eds., 2005); Bermann, *supra* note 2, at 73.

<sup>37</sup> See Gilles Cuniberti, *Beyond Contract—The Case for Default Arbitration in International Commercial Disputes*, 32 *FORDHAM INT’L L.J.* 417, 418 (2009) (“Some of the major international arbitral institutions report that their caseload has increased dramatically.”); *The AAA/ICDR and Fidal’s “Dispute-Wise” Business Management France Survey Results Released*, 4 *ICDR INT’L ARB. REP.* 3, 3 (2013) (identifying administration of 996 cases during 2012); International Chamber of Commerce, *ICC Reveals Record Number of New Arbitration Cases Filed in 2016*, <https://iccwbo.org/media-wall/news-speeches/icc-reveals-record-number-new-arbitration-cases-filed-2016/> (last visited Mar. 28, 2017) (identifying 966 arbitration requests filed at the ICC in 2016); LCIA, *REGISTRAR’S REPORT 1* (2013), <http://www.lcia.org/LCIA/reports.aspx> (identifying 290 arbitrations filed in 2013); *SCC Statistics 2014*, *ARBITRATION INST. STOCKHOLM CHAMBER OF COMMERCE*, <http://www.sccinstitute.com/media/93526/statistics-2014.pdf> (identifying 117 new arbitrations in 2013).

<sup>38</sup> See Mark Bezan, James Nicholson & Howard Rosen, *Trends in International Arbitration: A New World Order*, *FTI JOURNAL* 3–4 (Feb. 2015), [http://www.ftijournal.com/uploads/images/GAR\\_020415.pdf](http://www.ftijournal.com/uploads/images/GAR_020415.pdf) (reporting that in 2012, there were over 2700 international arbitration cases filed in various institutions and, in 2013, the value of pending “major claims” was over US\$1.6 billion); Richard W. Naimark & Stephanie E. Keer, *Analysis of UNCITRAL Questionnaires on Interim Relief*, in *TOWARDS A SCIENCE OF INTERNATIONAL ARBITRATION: COLLECTED EMPIRICAL RESEARCH* 129, 129 (Christopher R. Drahozal & Richard W. Naimark eds., 2005) (observing that, in 2000, the AAA administered over 500 disputes, and over the years, the ICC has administered cases “with claims in the billions of dollars”).

<sup>39</sup> Michael D. Goldhaber, *Arbitration Scorecard 2015*, *AM. LAWYER: FOCUS EUROPE* (2015). Goldhaber’s article does not indicate, however, whether the cases he analyzed were filed, pending, or just randomly sampled during the time of analysis (2013–2014).

Early arbitrations addressed disputes under the 1794 Jay Treaty,<sup>40</sup> involving claims about wartime debts owed to British merchants and which owes its origin to the advocacy of Alexander Hamilton,<sup>41</sup> and the *Alabama Claims Commission*, where arbitrators adjudicated disputes involving destroyed U.S. commercial vessels.<sup>42</sup>

ITA is the method of dispute resolution embedded in more than 3000 bilateral and multilateral investment treaties. These treaties grant foreign investors substantive rights and provide *ex ante* consent to arbitration.<sup>43</sup> Parties resolve disputes arising under the treaties, including claims of improper discrimination, failure to provide proper compensation for expropriation, and breaches of promises to provide “fair and equitable” treatment.<sup>44</sup> Only qualifying investors can sue, and they can only sue for state conduct breaching a treaty causing compensable damage.<sup>45</sup> ITA disputes receiving public attention include: investors suing Argentina for damages after Argentina devalued the peso and other emergency measures to stabilize its economy;<sup>46</sup> the Chevron–Ecuador dispute over activities in the Amazonian rain forest<sup>47</sup> where the U.S. Supreme Court recently left in place a D.C. Circuit opinion confirming a US\$96

---

<sup>40</sup> Treaty of Amity, Commerce and Navigation, U.S.–Gr. Brit., Nov. 19, 1794, 8 Stat. 116; Richard B. Lillich, *The Jay Treaty Commissions*, 37 ST. JOHN’S L. REV. 260, 261–62 (1963).

<sup>41</sup> See RON CHERNOW, ALEXANDER HAMILTON 485–503 (2004) (discussing historical aspects of Jay Treaty); TODD ESTES, THE JAY TREATY DEBATE, PUBLIC OPINION, AND THE EVOLUTION OF EARLY AMERICAN POLITICAL CULTURE 82–83 (Sidney M. Milkis & Jerome M. Mileur eds., 2006) (describing initial reluctance by Hamilton and but noting his vigorous defense and support of the treaty).

<sup>42</sup> Charles H. Brower, II, *The Functions and Limits of Arbitration and Judicial Settlement Under Private and Public International Law*, 18 DUKE J. COMP. & INT’L L. 259, 272–74 (2008); Barton Legum, *The Innovation of Investor-State Arbitration under NAFTA*, 43 HARV. INT’L L.J. 531, 536 (2002).

<sup>43</sup> Howard Mann, *Reconceptualizing International Investment Law: Its Role in Sustainable Development*, 17 LEWIS & CLARK L. REV. 521, 523–24 (2013).

<sup>44</sup> Some rights are analogous to a constitutional “bill of rights” for investors. Susan D. Franck, *The Nature and Enforcement of Investor Rights Under Investment Treaties: Do Investment Treaties Have a Bright Future?*, 12 U.C. DAVIS J. INT’L L. & POL’Y 47, 48 (2005); David Schneiderman, *Investment Rules and the New Constitutionalism*, 25 LAW & SOC. INQUIRY 757, 767 (2000). States can limit court access with sovereign immunity, fail to permit domestic review of government conduct, or have strong rule of law. Stephen E. Blythe, *The Advantages of Investor-State Arbitration as a Dispute Resolution Mechanism in Bilateral Investment Treaties*, 47 INT’L LAW. 273, 274–75, 281–82 (2013).

<sup>45</sup> Susan D. Franck & Lindsey Wylie, *Predicting Outcomes in Investment Treaty Arbitration*, 65 DUKE L.J. 459, 473–74 (2015).

<sup>46</sup> JOSÉ E. ALVAREZ, THE PUBLIC INTERNATIONAL LAW REGIME GOVERNING INTERNATIONAL INVESTMENT 248 (2011).

<sup>47</sup> *Chevron Corp. v. Ecuador*, PCA Case No. 2009-23, Fourth Interim Award on Interim Measures (Perm. Ct. Arb. 2013), <http://www.italaw.com/sites/default/files/case-documents/italaw1274.pdf>; Jesse Greenspan, *2nd Circ. Greenlights Chevron, Ecuador Arbitration*, LAW360 (Mar. 17, 2011, 2:42 PM), <https://www.law360.com/articles/232959/2nd-circ-greenlights-chevron-ecuador-arbitration> (explaining that the dispute is about “pollution in the Amazon rain forest”).

million award,<sup>48</sup> and Phillip Morris suing Australia and Uruguay (and losing both cases) for plain-packaging cigarette regulations that arguably resulted in expropriation of intellectual property.<sup>49</sup> Other ITA disputes are less sensational and more business-oriented, including suits involving revocation of a banking license or failure to pay dividends.<sup>50</sup> ITA requires using applicable law, which is usually derived from the treaty, to make decisions.

Since the first award in 1990,<sup>51</sup> ITA has expanded. While the global ITA caseload is smaller than ICA,<sup>52</sup> the value at stake is nonetheless noteworthy. The average ITA claim exceeds US\$650 million, the average combined legal fees are roughly US\$10 million, and arbitrators and institutional expenses cost roughly US\$1 million per case.<sup>53</sup> Experts have estimated that ITA will cover roughly 40%–60% of global investment.<sup>54</sup>

---

<sup>48</sup> *Chevron Corp. v. Ecuador*, 795 F.3d 200, 203, 209 (D.C. Cir. 2015), *cert denied*, 136 S. Ct. 2410 (2016); Caroline Simson, *A Cheat Sheet to Chevron's Epic Feud with Ecuador*, LAW360 (June 14, 2016), <https://www.law360.com/articles/805987/a-cheat-sheet-to-chevron-s-epic-feud-with-ecuador>.

<sup>49</sup> *Philip Morris Asia Ltd. v. Australia*, PCA Case No. 2012-12, Award on Jurisdiction and Admissibility, (Perm. Ct. Arb. 2015), <http://www.pccases.com/web/sendAttach/1711>; *Philip Morris Brand Sàrl v. Oriental Republic of Uruguay*, ICSID ARB/10/7, Award (July 8, 2016), <http://www.italaw.com/sites/default/files/case-documents/italaw7417.pdf>.

<sup>50</sup> See Ross P. Buckley & Paul Blyschak, *Guarding the Open Door: Non-party Participation Before the International Centre for Settlement of Investment Disputes*, 22 *BANKING & FIN. L. REV.* 353, 366 (2007); Michael Waibel, *Opening Pandora's Box: Sovereign Bonds in International Arbitration*, 101 *AM. J. INT'L L.* 711, 748 (2007).

<sup>51</sup> *Asian Agric. Prods. Ltd. v. Republic of Sri Lanka*, ICSID Case No. ARB/87/3, Final Award (June 27, 1990), <http://www.italaw.com/sites/default/files/case-%20documents/ita1034.pdf>.

<sup>52</sup> The United Nations Conference on Trade and Development estimates there have been roughly 500 ITA disputes. Susan D. Franck, *Conflating Politics and Development? Examining Investment Treaty Arbitration Outcomes*, 55 *VA. J. INT'L L.* 13, 15 (2014); see also Bezzant, Nicholson & Rosen, *supra* note 38, at 3 (estimating roughly 40–50 ICSID cases are filed every year).

<sup>53</sup> Franck & Wylie, *supra* note 45.

<sup>54</sup> *Trans-Pacific Partnership: Summary of U.S. Objectives*, OFFICE OF THE U.S. TRADE REPRESENTATIVE, <https://ustr.gov/tpp/Summary-of-US-objectives> (last visited Feb. 1, 2017); Jana Kasperkevic, *You Down with TPP? An Explainer on Obama's 'Secret' Trade Pact*, THE GUARDIAN (May 12, 2015, 10:10 PM), <https://www.theguardian.com/us-news/2015/may/12/trans-pacific-partnership-explainer>; Rem Korteweg, *It's the Geopolitics, Stupid: Why TTIP Matters*, CTR. FOR EUROPEAN REFORM (Apr. 2, 2015), <http://www.cer.org.uk/insights/it%E2%80%99s-geopolitics-stupid-why-ttip-matters>. Should TPP not go into effect, the estimate would require reconsideration. In any event, the U.S. withdrawal from TPP may require recalculation. See *supra* note 15 (indicating that the future and scope of TPP is uncertain).

Beyond sheer caseload and fiscal risks, international arbitration regulates global economic activity<sup>55</sup> and contributes to transnational lawmaking.<sup>56</sup> In ICA, tribunals render a final, binding, and enforceable decision for disputes arising under national law. These decisions create law for courts supervising arbitration or evaluating award enforceability. In ITA, tribunals evaluate treaty obligations to ascertain whether state conduct violates an investor's treaty-protected rights. This requires assessment of state liability for international law wrongs and can involve public policy considerations.<sup>57</sup> Although neither ICA nor ITA necessarily creates *de jure* precedent,<sup>58</sup> arbitration awards have a *de facto* effect<sup>59</sup> and have the capacity to influence doctrinal development.<sup>60</sup> ICA and ITA lack a unified traditional court structure, but the arbitral mandate requires arbitrators to apply law to facts and to render binding decisions that can be reviewed by national courts.<sup>61</sup>

---

<sup>55</sup> See Barbara Koremenos, *If Only Half of International Agreements Have Dispute Resolution Provisions, Which Half Needs Explaining?*, 36 J. LEGAL STUD. 189, 190 (2007); W. Michael Reisman, *International Investment Arbitration and ADR: Married but Best Living Apart*, 24 ICSID REV.—FOREIGN INV. L.J. 185, 186, 189 (2009); Jeswald W. Salacuse, *Is There a Better Way? Alternative Methods of Treaty-Based, Investor-State Dispute Resolution*, 31 FORDHAM INT'L L.J. 138, 138–39 (2007).

<sup>56</sup> Thomas E. Carbonneau, *Judicial Approval in Building the Civilization of Arbitration*, 113 PENN. ST. L. REV. 1343, 1344 (2009); Stephan W. Schill, *International Arbitrators as System-Builders*, 106 AM. SOC'Y INT'L L. PROC. 295, 295 (2012).

<sup>57</sup> Julie A. Maupin, *Public and Private in International Investment Law: An Integrated Systems Approach*, 54 VA. J. INT'L L. 367, 370–78 (2014).

<sup>58</sup> Franck, *supra* note 4, at 1611–12; see also Rogers, *supra* note 3, at 999–1000 (“[I]n the absence of a formal system of stare decisis, and despite the confidential and ‘private’ nature of international arbitration, arbitration proceedings generate procedural rules and practices, and to a lesser extent substantive rules, that serve as precedent for future arbitrations and beyond.”); *Id.* at 999 n.145 (“[P]ublished awards fail to ‘command stare decisis respect’ like a court decision[.]”).

<sup>59</sup> Jason Webb Yackee, *Controlling the International Investment Law Agency*, 53 HARV. INT'L L.J. 391, 413 (2012); Strong, *supra* note 25, at 504.

<sup>60</sup> W. Mark C. Weidemaier, *Toward a Theory of Precedent in Arbitration*, 51 WM. & MARY L. REV. 1895, 1929 (2010).

<sup>61</sup> International arbitration falls squarely within the ambit of international courts and tribunals. Gary B. Born, *A New Generation of International Adjudication*, 61 DUKE L.J. 775, 780–81 (2012); Andrea K. Bjorklund, *Private Rights and Public International Law: Why Competition Among International Economic Law Tribunals Is Not Working*, 59 HASTINGS L.J. 241, 245 (2007); Lucy Reed, *Great Expectations: Where Does the Proliferation of International Dispute Resolution Tribunals Leave International Law?*, 96 AM. SOC'Y INT'L L. PROC. 219 (2002). In ICA, the New York Convention permits limited review of arbitration awards by national courts. In ITA, disputes rendered pursuant to the New York Convention are similarly reviewable by national courts, whereas disputes rendered under the ICSID Convention benefit from internal annulment proceedings but are only subject to review by national courts as if the award was a national court judgment. Franck, *Legitimacy Crisis*, *supra* note 4, at 1546–55.

## B. Arbitration Procedures

International arbitration is a creature of consent. Parties—whether individuals, commercial entities, or governments—must agree to arbitrate conflicts involving commercial disputes or other transnational relationships.<sup>62</sup> In ICA, parties typically agree to arbitrate in contracts *ex ante*,<sup>63</sup> and in ITA, two or more states make an *ex ante* offer in a treaty that their respective investors' can arbitrate, which investors later accept by initiating arbitration.<sup>64</sup> Under both ICA and ITA, parties agree arbitrators will be neutral adjudicators finally resolving disputes using applicable law. Arbitration allows parties to create tailor-made procedural rules, but practically speaking, particularly with its “judicialization”<sup>65</sup> or “Americanization,”<sup>66</sup> international arbitration procedures resemble more rigid and rule-oriented national court litigation.<sup>67</sup> International arbitration frequently involves submission of formal pleadings (e.g., a Request for Arbitration and Answer), requests to dismiss cases early on jurisdictional grounds, petitions for interim relief, requests for documents; competing expert reports, hearings for the evidence presentation and oral testimony, witness cross-examination, post-hearing submissions, and formal awards.<sup>68</sup>

---

<sup>62</sup> See, e.g., BORN, *supra* note 11, at 187, 197, 217; LEW, MISTELIS & KROLL, *supra* note 33, at 4–5, 99–186.

<sup>63</sup> See Alan Scott Rau & Edward F. Sherman, *Tradition and Innovation in International Arbitration Procedure*, 30 TEX. INT'L L.J. 89, 90 (1995).

<sup>64</sup> See Anna T. Katselas, *Exit, Voice, and Loyalty in Investment Treaty Arbitration*, 93 NEB. L. REV. 313, 314 (2014); Jan Paulsson, *Arbitration Without Privity*, 10 ICSID REV. FOREIGN INV. L.J. 232, 233 (1995).

<sup>65</sup> INTERNATIONAL ARBITRATION IN THE 21ST CENTURY: TOWARDS “JUDICIALIZATION” AND UNIFORMITY?, at ix (Richard B. Lillich & Charles N. Brower eds., 1994); Winston Stromberg, *Avoiding the Full Court Press: International Commercial Arbitration and Other Global Alternative Dispute Resolution Processes*, 40 LOY. L.A. L. REV. 1337, 1342 n.22 (2007).

<sup>66</sup> See, e.g., Roger P. Alford, *The American Influence on International Arbitration*, 19 OHIO ST. J. DISP. RESOL. 69, 69 (2003); Bernard Audit, *L'Américanisation du droit [The Americanization of Law]*, 45 ARCHIVES DE PHILOSOPHIE DU DROIT 7 (2001); Eric Bergsten, *The Americanization of International Arbitration*, 18 PACE INT'L L. REV. 289 (2006); Kenneth F. Dunham, *International Arbitration Is Not Your Father's Oldsmobile*, 2005 J. DISP. RESOL. 323, 326–27; Susan L. Karamanian, *Overstating the “Americanization” of International Arbitration: Lessons from ICSID*, 19 OHIO ST. J. DISP. RESOL. 5, 5–7 (2003); George M. von Mehren & Alana C. Jochum, *Is International Arbitration Becoming Too American?*, 2 GLOBAL BUS. L. REV. 47, 47–57 (2011).

<sup>67</sup> BORN, *supra* note 11, at 1–2; see also William W. Park, *Arbitrators and Accuracy*, 1 J. INT'L DISP. SETTLEMENT 25, 26–27 (2010) (“In examining the competing views of reality proposed by each side, arbitrators aim to get as near as reasonably possible to a correct picture of those disputed events, words, and legal norms that bear consequences for the litigants' claims and defences.”).

<sup>68</sup> See, e.g., BORN, *supra* note 11, at 1–2; Franck, *supra* note 20, at 192–94.

While parties control (either directly or indirectly) arbitrator appointment,<sup>69</sup> arbitrators must abide by rules that require impartial and independent decisionmaking.<sup>70</sup> The applicable law imposes duties upon arbitrators,<sup>71</sup> like minimizing expense and delay in decisionmaking.<sup>72</sup> After arbitrators render an award, treaties facilitate streamlined enforcement of awards.<sup>73</sup>

### C. Arbitration Decisionmaking

Given their mandate and discretion on issues of economic and doctrinal importance, international arbitrators play a vital role in global disputes. The integrity and quality of their decisionmaking is therefore central to arbitration's legitimacy as a form of dispute settlement.<sup>74</sup> Uncertainty about the quality of decisions has created apprehension and debate<sup>75</sup> about whether international arbitrators should be stripped of jurisdiction in favor of judges.<sup>76</sup>

---

<sup>69</sup> There are various appointment methods. Depending on parties' agreement and applicable law, parties, co-arbitrators, arbitral institutions, or another neutral body may appoint arbitrators; national courts can also make appointments. See, e.g., Born, *supra* note 11, at 614–52.

<sup>70</sup> See Chiara Giorgetti, *Who Decides Who Decides in International Investment Arbitration*, 35 U. PA. J. INT'L L. 431, 438–54 (2013); Franck, *supra* note 35, at 502–12; see also Craig, *supra* note 35, at 253 (“It is widely recognized in the practice of international commercial arbitration and in the rules of international arbitration institutions that a party-appointed arbitrator must be impartial and independent.” (footnote omitted)).

<sup>71</sup> LEW, MISTELIS & KRÖLL, *supra* note 33, at 279–83; Cindy G. Buys, *The Arbitrators' Duty to Respect the Parties' Choice of Law in Commercial Arbitration*, 79 ST. JOHN'S L. REV. 59, 59 (2005); Susan D. Franck, *The Liability of International Arbitrators*, 20 N.Y. L. SCH. J. INT'L & COMP. L. 1, 4–11, 37, 44 (2000).

<sup>72</sup> England and Wales impose an obligation to “adopt procedures suitable to the circumstances of the particular case, avoiding unnecessary delay or expense, so as to provide a fair means for the resolution” of disputes. Arbitration Act 1996, c. 23, § 33(1) (Eng.), <http://www.legislation.gov.uk/ukpga/1996/23/contents>.

<sup>73</sup> Convention on the Recognition and Enforcement of Foreign Arbitral Awards art. 1, June 10, 1958, 21 U.S.T. 2517, 330 U.N.T.S. 38; Inter-American Convention on International Commercial Arbitration, Jan. 30, 1975, O.A.S.T.S. No. 42; Convention on the Settlement of Investment Disputes Between States and Nationals of Other States art. 53, Mar. 18, 1965, 17 U.S.T. 1270, 575 U.N.T.S. 159.

<sup>74</sup> Stephan W. Schill, *International Arbitrators as System-Builders*, 106 AM. SOC'Y INT'L L. PROC. 295, 296–97 (2012).

<sup>75</sup> Patrick Sweeney, *Exceeding Their Powers: A Critique of Stolt-Nielsen and Manifest Disregard, and a Proposal for Substantive Arbitral Award Review*, 71 WASH. & LEE L. REV. 1571, 1574 (2014).

<sup>76</sup> See Editorial, *The Arbitration Game*, THE ECONOMIST (Oct. 11, 2014), <http://www.economist.com/news/finance-and-economics/21623756-governments-are-souring-treaties-protect-foreign-investors-arbitration>; Henry Farrell, *People Are Freaking Out About the Trans Pacific Partnership's Investor Dispute Settlement System. Why Should You Care?*, WASH. POST (Mar. 26, 2015), <https://www.washingtonpost.com/news/monkey-cage/wp/2015/03/26/people-are-freaking-out-about-the-trans-pacific-partnerships-investor-dispute-settlement-system-why-should-you-care/>; Jonathan Weisman, *Trans-Pacific Partnership Seen as Door for Foreign Suits Against U.S.*, N.Y. TIMES (Mar. 25, 2015), [https://www.nytimes.com/2015/03/26/business/trans-pacific-partnership-seen-as-door-for-foreign-suits-against-us.html?\\_r=2](https://www.nytimes.com/2015/03/26/business/trans-pacific-partnership-seen-as-door-for-foreign-suits-against-us.html?_r=2); *supra* notes 14–17; see also Juergen Mark, *German Association of Judges on the TTIP Proposal of the European Commission*, GLOBAL ARB. NEWS (Mar. 21, 2016), <https://globalarbitrationnews.com/german-association-judges-proposal-european-commission-introduction-investment-court-system-settle-investor-state-disputes-transatlantic-trade-investment/>

Researchers have studied judges and found that they do not decide cases in a purely rational manner. Instead, judges often make initial intuitive judgments which they might, or might not, override with deliberation.<sup>77</sup> They are influenced, for example, by irrelevant numerical anchors,<sup>78</sup> the way outcomes are framed,<sup>79</sup> and irrelevant emotional cues.<sup>80</sup>

In stark contrast to this research on judges, we know little about arbitrators. Some researchers have conducted qualitative arbitrator interviews<sup>81</sup> or examined

---

(describing how German judges reject any form of transnational dispute settlement involving suits against states and instead assert national court judges should have jurisdiction); *TTIP Trade Talks: German Judges Oppose New Investor Courts*, BBC NEWS (Feb. 5, 2016), <http://www.bbc.com/news/world-europe-35503885> (same).

<sup>77</sup> See generally Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Blinking on the Bench: How Judges Decide Cases*, 93 CORNELL L. REV. 1 (2007) [hereinafter Guthrie, Rachlinski & Wistrich, *Blinking*]; Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *The "Hidden Judiciary": An Empirical Examination of Executive Branch Justice*, 58 DUKE L.J. 1477 (2009) [hereinafter Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*]; Guthrie, Rachlinski & Wistrich, *supra* note 10; Andrew J. Wistrich, Chris Guthrie & Jeffrey J. Rachlinski, *Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding*, 153 U. PENN. L. REV. 1251 (2005) [hereinafter Wistrich, Guthrie & Rachlinski, *Disregarding*]. The theories of either a pure formalist or pure realist model of decisionmaking are unsupported by the data; rather the data supports a model of judging called the "intuitive override" model, whereby adjudication involves initial intuitive assessments that can be tested against evidence and logic. See, e.g., Guthrie, Rachlinski & Wistrich, *Blinking, supra*; Linda A. Berger, *A Revised View of the Judicial Hunch*, 10 LEGAL COMM. & RHETORIC: JALWD 1, 17–18 (2013).

<sup>78</sup> See *infra* notes 143–51; see also Jeffrey J. Rachlinski, Andrew J. Wistrich, & Chris Guthrie, *Can Judges Make Reliable Numeric Judgments? Distorted Damages and Skewed Sentences*, 90 IND. L.J. 695 (2015) [hereinafter Rachlinski, Wistrich & Guthrie, *Distorted Damages*].

<sup>79</sup> See *infra* notes 187–92 (discussing framing).

<sup>80</sup> See Jeffrey J. Rachlinski, Andrew J. Wistrich & Chris Guthrie, *Altering Attention in Adjudication*, 60 UCLA L. REV. 1586 (2013) (identifying how directing judicial attention shapes outcomes); Jeffrey J. Rachlinski, Chris Guthrie & Andrew J. Wistrich, *Contrition in the Courtroom: Do Apologies Affect Adjudication?*, 98 CORNELL L. REV. 1189 (2013) [hereinafter Rachlinski, Guthrie & Wistrich, *Contrition*] (finding apologies can induce judges to be more lenient but identifying the limitations of apologies); Andrew J. Wistrich, Jeffrey J. Rachlinski & Chris Guthrie, *Heart Versus Head: Do Judges Follow the Law or Follow Their Feelings?*, 93 TEX. L. REV. 855, 862 (2015) [hereinafter Wistrich, Rachlinski & Guthrie, *Heart*] (finding that "judges' feelings about litigants influence their judgments").

<sup>81</sup> See, e.g., YVES DEZALAY & BRYANT G. GARTH, *DEALING IN VIRTUE: INTERNATIONAL COMMERCIAL ARBITRATION AND THE CONSTRUCTION OF A TRANSNATIONAL LEGAL ORDER* (1996); Thomas Schultz & Robert Kovacs, *The Rise of a Third Generation of Arbitrators? Fifteen Years After Dezalay & Garth*, 28 ARB. INT'L 161 (2012); see also JOSHUA KARTON, *THE CULTURE OF INTERNATIONAL ARBITRATION AND THE EVOLUTION OF CONTRACT LAW* 10 (2013) (drawing upon interviews with international arbitrators "selected to represent as wide as possible a range of backgrounds" to conclude arbitrators and judges decide cases differently); Sophie Nappert & Dieter Flader, *Psychological Factors in the Arbitral Process*, in *THE ART OF ADVOCACY IN INTERNATIONAL ARBITRATION* 134 (Doak Bishop & Edward G. Kehoe eds., 2d ed. 2010) (exploring "what persuades and triggers decision-making in international arbitrators" by circulating questionnaires on listservs, receiving nineteen responses, and failing to identify a response rate).



arbitration outcomes.<sup>82</sup> We are unaware of any scholarship experimentally testing international arbitrator decisionmaking.<sup>83</sup> Until now.

## II. RESEARCH QUESTIONS AND METHODOLOGY

In this Part, we introduce our study of arbitrator decisionmaking, including our research hypotheses, the demographic characteristics of our participants, and our experimental methodology.

### A. *Research Hypotheses*

It is an open question whether international arbitrators, like other experts, are influenced by cognitive illusions. Given the existing experimental literature on national court and administrative law judges (ALJs), one reasonable theory is that, like other adjudicators, cognitive illusions affect international arbitrators. An alternative theory is that cognitive illusions affect international arbitrators differently, presumably making them inferior adjudicators and thereby less worthy of resolving complex international disputes.

---

<sup>82</sup> See, e.g., Christopher R. Drahozal, *Behavioral Analysis of Arbitral Decision Making*, in TOWARDS A SCIENCE OF INTERNATIONAL ARBITRATION: COLLECTED EMPIRICAL RESEARCH 319 (Christopher R. Drahozal & Richard W. Naimark eds., 2005) (exploring ICA empirical literature); Susan D. Franck, *Development and Outcomes in Investment Treaty Arbitration*, 50 HARV. INT'L L.J. 435, 438 (2009) (exploring whether the context, political or otherwise, of arbitration explains ITA outcomes); Franck & Wylie, *supra* note 45 (exploring arbitrator-based and case-based models of ITA outcomes); Daphna Kapeliuk, *The Repeat Appointment Factor: Exploring Decision Patterns of Elite Investment Arbitrators*, 96 CORNELL L. REV. 47 (2010) (exploring appointment patterns on arbitration outcomes); see also Sergio Puig, *Social Capital in the Arbitration Marketplace*, 25 EUROPEAN J. INT'L L. 387 (2014) (exploring the web of the arbitrator marketplace in ITA).

<sup>83</sup> In 2004, Drahozal observed, “[e]mpirical studies of the prevalence of cognitive illusions in arbitral decisionmaking are exceedingly rare. I am aware of no such studies using experimental techniques.” Christopher R. Drahozal, *A Behavioral Analysis of Private Judging*, 67 LAW & CONTEMP. PROBS. 105, 114 (2004). This remains true in international arbitration. Scholars, like Drahozal, have largely explored the theoretical application of cognitive illusions to international dispute settlement. See, e.g., Shari Seidman Diamond, *Psychological Aspects of Dispute Resolution: Issues for International Arbitration*, in INTERNATIONAL COUNCIL FOR COMMERCIAL ARBITRATION: IMPORTANT CONTEMPORARY QUESTIONS 327 (Albert Jan van den Berg ed., 2003); Jan-Philip Elm, *Behavioral Insights into International Arbitration: An Analysis of How to De-Bias Arbitrators*, 27 AM. REV. INT'L ARB. 74 (2016); Ernest A. Haggard & Soia Mentschikoff, *Responsible Decision Making in Dispute Settlement*, in LAW, JUSTICE, AND THE INDIVIDUAL IN SOCIETY: PSYCHOLOGICAL AND LEGAL ISSUES 277 (June Louin Tapp & Felice J. Levine eds., 1977); Lucy Reed, *The 2013 Hong Kong International Arbitration Centre Kaplan Lecture—Arbitral Decision-Making: Art, Science or Sport?*, 30 J. INT'L ARB. 85 (2013); Edna Sussman, *Arbitrator Decision Making: Unconscious Psychological Influences and What You Can Do About Them?*, 24 AM. REV. INT'L ARB. 487 (2013). A study published after this article was accepted for publication experimentally explores the cognitive illusions of a small group of domestic arbitrators. Rebecca Helm, Andrew J. Wistrich & Jeffrey J. Rachlinski, *Are Arbitrators Human?*, 13 J. EMPIRICAL LEG. STUD. 666 (2016).

Given existing research on judges, we began our study with five descriptive research hypotheses designed to shed light on the extent to which arbitrators make intuitive, impressionistic decisions or deliberative and fully rational decisions:

1. International arbitrators solve generic problems in an intuitive, rather than deliberative, manner.
2. When faced with a concrete international dispute, international arbitrators are influenced by relevant and irrelevant numeric anchors when awarding damages.
3. International arbitrators respond more strongly to the possibility of losses and less strongly to the possibility of gains when deciding disputes.
4. International arbitrators resolve disputes based on representative cues rather than deliberative reason.
5. International arbitrators are prone to egocentric bias when evaluating themselves and the disputes they address.

We also sought, where possible, to compare arbitrators to national judges who responded to similar hypothetical vignettes in earlier research. Because we did not provide the same problems to judges and arbitrators, and because we did not test them at the same time, we are limited in our ability to make statistically valid comparisons. Where comparison seemed legitimate, we hypothesized that judges would outperform arbitrators, rendering more deliberative decisions.<sup>84</sup>

#### *B. International Arbitrators: Participants*

Our target population was international arbitrators. Unlike national judges, there is no unified repository identifying all individuals willing to serve, or with a history of serving, as an international arbitrator. This is, in part, because the international arbitration community changes frequently and unpredictably as people enter and exit the profession.<sup>85</sup>

We sampled arbitrators attending the prestigious biennial Congress of the International Council for Commercial Arbitration (ICCA) in 2014 in Miami. These participants had no special interest in psychology or psychological

---

<sup>84</sup> Given the elite and competitive international arbitration market, our research hypothesis could have been that arbitrators exhibit superior cognition. Null-Hypothesis Significance Testing tests both hypotheses, as the objective is to identify group differences.

<sup>85</sup> See DEZALAY & GARTH, *supra* note 81, at 12, 28, 61, 117, 157, 242, 248, 296 (1996); Catherine A. Rogers, *Gulliver's Troubled Travels, or the Conundrum of Comparative Law*, 67 GEO. WASH. L. REV. 149, 167 (1998).

research.<sup>86</sup> ICCA is an important group in the international arbitration community, and the biennial ICCA Congress is a prominent event that many international arbitrators attend. ICCA therefore provided a singular opportunity to research international arbitrators.<sup>87</sup>

At the 2014 ICCA conference, 1031 professionals registered for the Congress.<sup>88</sup> Based on cross-referencing registered participants with publicly available information,<sup>89</sup> we identified 496 registrants (roughly 48% of attendees) with experience as an international arbitrator.

We administered materials to all registrants attending the first plenary session. After excluding four individuals requesting their responses be omitted from published research, 548 international arbitration specialists completed the experiment. As our hypotheses involved the cognition of international arbitrators, rather than counsel,<sup>90</sup> we excluded responses from subjects who had never acted as an arbitrator.<sup>91</sup> We therefore analyzed responses from 262 individuals who self-identified as having been an arbitrator<sup>92</sup> in at least one ICA or ITA dispute.<sup>93</sup> These participants represented roughly 48% of all registrants

---

<sup>86</sup> ICCA is a prestigious non-governmental organization of the international arbitration bar. ICCA's governing board includes prominent arbitrators, the ICSID secretary general, past presidents of the American Society of International Law, Principal Legal Counsel for the Government of Mexico in negotiating NAFTA, General Counsel of ExxonMobil, Attorney General of Kenya, Pakistan's former Attorney General, Singapore's Chief Justice of the Supreme Court, Chair of the Hong Kong International Arbitration Centre, Director of the Cairo Regional Centre for International Commercial Arbitration, and authors of several core international arbitration treatises. Franck et al., *supra* note 4, at 441; Franck, et al., *International Arbitration: Demographics, Precision and Justice*, in LEGITIMACY: MYTHS, REALITIES, CHALLENGES, ICCA CONGRESS SERIES NO. 18, at 33, 57–9 [hereinafter Franck et al., ICCA MIAMI CONGRESS PROCEEDINGS]. None of the authors are ICCA members.

<sup>87</sup> Franck et al., *supra* note 4, at 440–42.

<sup>88</sup> See Franck et al., *supra* note 4, at 441 & n.35 (noting, as twelve registrants worked on the research team and two people reviewed earlier drafts, “only 1,017 of the registrants were capable of answering the survey”). ICCA Congress Proceedings reflect the large, transnational attendees. *List of Participants*, in LEGITIMACY: MYTHS, REALITIES, CHALLENGES, ICCA CONGRESS SERIES NO. 18, at 1041 (Albert Jan van den Berg ed., 2015).

<sup>89</sup> We cross-referenced attendee lists with past arbitrator activity in *Who's Who Legal*, *Chambers & Partners*, IAI Paris, *Global Arbitration Review*, company websites, and Google searches. Special thanks is owed to Stephanie Miller, a research librarian at Washington & Lee University School of Law where the lead author formerly worked, for undertaking this background research.

<sup>90</sup> Future research might explore counsel, or the cognition of others in international arbitration, including insurers, third-party funders, experts, parties, or policy makers.

<sup>91</sup> When analyzing those serving as counsel, results tended to be similar. A full discussion of variations between counsel and arbitrators is beyond the scope of this Article. As arbitrators serve as counsel—and international arbitrators are drawn from the arbitration bar—similarities would be unsurprising.

<sup>92</sup> By walking up and down the rows in a large conference hall, we visually observed that many of the subjects completing the survey were arbitrators. Franck et al., *supra* note 4, at 443.

<sup>93</sup> Some participants failed to state they were ICA or ITA arbitrators. *Id.* at 448 n.57.

returning materials and 53% of arbitrators registered for the ICCA Congress.<sup>94</sup> Sixty-seven ITA arbitrators responded. This represents a reasonable proportion (27%) of known ITA arbitrators.<sup>95</sup>

The arbitrator sample included 46 (17.6%) women and 216 (82.4%) men.<sup>96</sup> The average age of an arbitrator was 54,<sup>97</sup> and the average male arbitrator was reliably older than the average female.<sup>98</sup> Most of the international arbitrators were from a developed country.<sup>99</sup> Despite data collection in Miami, the largest proportion (48%) of arbitrators were European,<sup>100</sup> and English was the primary native language of 43.3%.<sup>101</sup> For legal training, 38.5% of arbitrators were exclusively trained in common law, 33.8% were exclusively trained in civil law, and 27.7% had training in *both* common and civil law. Arbitrators in our sample had decided thirty-five cases on average.<sup>102</sup>

### C. *Experimental Method*

We created stimulus materials to assess arbitrators' decisionmaking by asking them to resolve mock disputes using brief case vignettes. We created scenarios mirroring realistic international commercial and investment disputes and then used the arbitrators' responses to those scenarios to assess arbitration

---

<sup>94</sup> *Id.* at 443 n.44.

<sup>95</sup> See SUSAN D. FRANCK, MYTHS AND REALITIES IN INVESTMENT TREATY ARBITRATION (forthcoming) (coding ITA arbitrators on tribunals rendering public awards); Puig, *supra* note 82, at 403 (coding ICSID arbitrator appointments and identifying 419 arbitrators).

<sup>96</sup> Franck et al., *supra* note 4, at 453.

<sup>97</sup> *Id.*

<sup>98</sup> The mean age was 55.8 for male arbitrators, and 47.5 for female arbitrators. *Id.* at 453–55. The age difference was statistically significant and medium-sized. *Id.* at 454. The gender demographics and age breakdown have been replicated by research from practitioners. For example, the International Chamber of Commerce—one of the world's preeminent international arbitration institutions—recently identified that about 10% of their arbitrators were female, and female arbitrators were generally younger than male arbitrators. Mirèze Philippe, *Speeding Up the Path for Gender Equality*, 14 TRANSNAT'L DISP. MGMT., Jan. 2017, at 4; see also Lucy Greenwood & C. Mark Baker, *Is the Balance Getting Better? An Update on the Issue of Gender Diversity in International Arbitration*, 28 ARB. INT'L 413 (2015) (identifying historical gender balance issues in the field of arbitration and recent efforts, both internal and external, to redress the balance).

<sup>99</sup> This was true irrespective of whether “development status” derived from arbitrators' nationality using Organisation for Economic Co-operation and Development, World Bank, or United Nations Development Programme Human Development Index definitions. Franck et al., *supra* note 4, at 458–65.

<sup>100</sup> *Id.* at 459–60. Largest representation came from the United States (23.2%), United Kingdom (9.6%), France (8.8%), Brazil (7.2%), Switzerland (5.6%), Germany (4.8%), and Canada (4.8%). *Id.*

<sup>101</sup> *Id.* at 458–59. Other dominant primary languages were German (10.6%), French (10.2%), Portuguese (8.3%) and Spanish (7.1%). *Id.* Of the 205 participants fluent in a second language, 60.5% ( $n = 124$ ) spoke English, French 20.5% ( $n = 42$ ), Spanish 7.3% ( $n = 15$ ), and German 2% ( $n = 4$ ).

<sup>102</sup> The median was ten. The 25th and 75th percentile appointment levels were three and forty. *Id.* at 450.

decisionmaking. We followed protocols used in prior studies of judges—including U.S. state court judges,<sup>103</sup> U.S. federal court judges,<sup>104</sup> U.S. bankruptcy judges,<sup>105</sup> U.S. magistrates,<sup>106</sup> U.S. administrative law judges,<sup>107</sup> Canadian judges,<sup>108</sup> Dutch judges,<sup>109</sup> and Swiss judges.<sup>110</sup>

We presented a panel at the first plenary session entitled “Arbitration and Decision-Making: Live Empirical Study.”<sup>111</sup> The title was intentionally vague to avoid revealing our research topic before participants responded. At the beginning of the panel, we offered attendees the opportunity to complete a confidential survey. All attendees were orally instructed on protocols, including requests to read the survey, to take the materials seriously, to respond to each question in order, and to work independently without reference to others or internet searches.<sup>112</sup> We asked participants to complete the survey and instructed them not to identify themselves. We then distributed randomized surveys.

The survey materials began with a one-page instruction and consent form. The first page asked participants to read and respond to the questions independently and without discussing it with others, informed them that participation was voluntary and explained we intended to use responses during a follow-up presentation at the Congress.<sup>113</sup> The five remaining pages contained

---

<sup>103</sup> Guthrie, Rachlinski, & Wistrich, *Blinking*, *supra* note 77, at 10–11; Wistrich, Guthrie & Rachlinski, *Disregarding*, *supra* note 77, at 1279–82.

<sup>104</sup> Wistrich, Rachlinski & Guthrie, *Heart*, *supra* note 80, at 874–76; Wistrich, Guthrie & Rachlinski, *Disregarding*, *supra* note 77, at 1281–82.

<sup>105</sup> Jeffrey J. Rachlinski, Chris Guthrie & Andrew J. Wistrich, *Inside the Bankruptcy Judge's Mind*, 86 B.U. L. REV. 1227, 1230–32 (2006); *see also* Rachlinski, Guthrie & Wistrich, *Contrition*, *supra* note 80, at 1208–09 (evaluating apologies and adjudication for bankruptcy judges).

<sup>106</sup> Guthrie, Rachlinski & Wistrich, *supra* note 10, at 786–77.

<sup>107</sup> Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1491–94.

<sup>108</sup> Wistrich, Rachlinski & Guthrie, *Heart*, *supra* note 80, at 874–76; Rachlinski, Wistrich & Guthrie, *Distorted Damages*, *supra* note 78, at 720.

<sup>109</sup> Rachlinski, Wistrich & Guthrie, *Distorted Damages*, *supra* note 78, at 726.

<sup>110</sup> Mark Schweizer, *Kognitive Täuschungen vor Gericht [Cognitive Illusions in Court]*, DISSERTATION ZÜRICH (2005), [http://www.decisions.ch/dissertation/diss\\_methode.html](http://www.decisions.ch/dissertation/diss_methode.html) (analyzing Swiss judges through mail surveys).

<sup>111</sup> *See* INTERNATIONAL COUNCIL FOR COMMERCIAL ARBITRATION, *ICCA Miami Congress 2014 Working Programme* (Apr. 6, 2014), [http://www.arbitration-icca.org/media/2/14334105310240/icca\\_website\\_schedule\\_03.27.14.pdf](http://www.arbitration-icca.org/media/2/14334105310240/icca_website_schedule_03.27.14.pdf).

<sup>112</sup> We provided instructions orally and on the first page. *See* International Council for Commercial Arbitration, *Monday Plenary—ICCA Miami Congress 2014*, ARBITRATION-ICCA.ORG (Apr. 7, 2014), [http://www.arbitration-icca.org/conferences-and-congresses/ICCA\\_MIAMI\\_2014-video-coverage/ICCA\\_MIAMI\\_2014\\_Plenary\\_Session\\_7\\_April.html](http://www.arbitration-icca.org/conferences-and-congresses/ICCA_MIAMI_2014-video-coverage/ICCA_MIAMI_2014_Plenary_Session_7_April.html) (36:52–43:22).

<sup>113</sup> Subjects had the option to avoid use of their data in published research; four participants exercised this option. *Cf.* Guthrie, Rachlinski & Wistrich, *supra* note 10, at 787 (noting one judge of 168 declined to have responses used).

the survey materials; four pages pertained to the experiment and the remaining page asked questions related to Congress themes of legitimacy and precision.<sup>114</sup> For the experimental materials, each participant received questions to test our hypotheses.<sup>115</sup> To create controlled experimental conditions, although neither the introductory instructions nor the first page indicated this, we devised multiple versions of several scenarios which were randomly assigned to participants. Subjects had approximately thirty-five minutes to complete the survey. During survey administration, participants remained in the room, kept silent, and appeared to take the process seriously.

All session attendees returned the survey—whether fully completed, partially completed, or left blank—before leaving the plenary session. In total, 98.2% of the attendees answered at least one question.

### III. RESULTS

We found evidence that arbitrators, like judges, tended to make intuitive decisions and were influenced by well-known cognitive illusions like anchoring, framing, and the like. Where comparisons with judges were possible, we were generally unable to reliably distinguish between the responses of arbitrators and judges, suggesting the two groups performed comparably.<sup>116</sup> We also found evidence that arbitrators, as a group, were unlikely to merely “split the baby” between claimants and respondents. Taken together, the findings cast doubt on the bona fides of the normative narrative that international arbitrators should be stripped of jurisdiction and replaced by judges due to cognitive predisposition.

#### A. *Testing the Intuitive-Override Model*

We hypothesized that international arbitrators, like their judicial counterparts, make decisions using an “intuitive-override” model<sup>117</sup> whereby arbitrators may initially make an intuitive assessment that they could ultimately

---

<sup>114</sup> Demographic information and survey questions involving Congress themes are described elsewhere. Franck et al., *ICCA MIAMI CONGRESS PROCEEDINGS*, *supra* note 86, at 57–60; Franck et al., *supra* note 4, at 440–45.

<sup>115</sup> We created the materials over two years and beta-tested them on law students in St. Gallen, Switzerland, and Lexington, Virginia.

<sup>116</sup> There were only two instances when international arbitrators outperformed judges, namely one test comparing Cognitive Reflection Test scores with one group of state court judges and the representativeness hypothetical. *See infra* notes 132–35, 226. The two times we identified a reliable difference, the practical significance of the difference was small. The evidence, as measured and analyzed in our studies, never demonstrated that the intuitive cognition of international arbitrators was inferior to judges.

<sup>117</sup> *See supra* note 77 (discussing the intuitive-override model of adjudication).

override using more rational and deliberative cognition. To test this hypothesis, we administered the Cognitive Reflection Test (CRT), a simple test of deliberative reasoning described below.

### 1. CRT

Economist Shane Fredrick developed the CRT to test whether decisionmaking involves dual processing<sup>118</sup> where subjects have “the ability or disposition to resist reporting the response that first comes to mind.”<sup>119</sup> The CRT asks three questions. For each question, there is an intuitive but incorrect answer, as well as a correct answer that is easy to discern upon reflection.

The first CRT question is: “A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?” The intuitive response, 10¢, is mathematically incorrect. If the bat costs US\$1 more than 10¢ (US\$1.10) and the ball is 10¢, the total cost is US\$1.20. The correct answer is 5¢, with a bat costing US\$1.05 and a ball costing 5¢.<sup>120</sup> The calculation is relatively easy, but the analysis requires deliberation to avoid generating inadvertent error.

The second CRT question is: “If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?”<sup>121</sup> The intuitive answer is 100, but this is wrong. Deliberation reveals that if five machines make five widgets in five minutes, then each machine makes a single widget in five minutes. With that base rate, one can calculate it takes five minutes for 100 machines to make 100 widgets.<sup>122</sup>

The final CRT question asks: “In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take the patch to cover half of the lake?”<sup>123</sup> The intuitive (and incorrect) answer is twenty-four days. Using slower cognition to override snap judgments reveals the correct answer is forty-seven days. If the rate of growth means the amount doubles every day, compounding means half the lake was covered the day before (i.e., day forty-seven, not day forty-eight).

---

<sup>118</sup> Daniel Kahneman & Shane Frederick, *Representativeness Revisited: Attribute Substitution in Intuitive Judgment*, in *HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT*, 49, 51–52 (Thomas Gilovich, Dale Griffin & Daniel Kahneman eds., 2002).

<sup>119</sup> Shane Frederick, *Cognitive Reflection and Decision Making*, 19 *J. ECON. PERSP.*, Fall 2005, at 25, 35.

<sup>120</sup> *Id.* at 27, 37.

<sup>121</sup> *Id.* at 26–27.

<sup>122</sup> Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 10–11.

<sup>123</sup> Frederick, *supra* note 119, at 27.

## 2. Arbitrators' Performance

We analyzed responses from the arbitrators who answered all three questions.<sup>124</sup> International arbitrators' average CRT score was 1.47,<sup>125</sup> which exceeds mean CRT scores of judges participating in prior studies as shown in Table 1.

Table 1: Overall CRT Results – Arbitrators Compared to Judges and Others (*n*)

Sample ( <i>n</i> )	Mean	0 Correct Answers	1 Correct Answer	2 Correct Answers	3 Correct Answers
MIT students (61)	2.18	7%	16%	30%	48%
Carnegie-Mellon students (746) <sup>126</sup>	1.51	25%	25%	25%	25%
International Arbitrators (239)	1.47	24.3% (58)	24.7% (59)	30.5% (73)	20.5% (49)
North American Lawyers (247) <sup>127</sup>	1.46	24.3%	26.3%	28.3%	21.1%
Administrative Law Judges (126) <sup>128</sup>	1.33	30.2%	27.8%	20.6%	21.4%
Florida Judges (252) <sup>129</sup>	1.23	30.6%	31.0%	23.8%	14.7%
University of Michigan: Ann Arbor students (1267) <sup>130</sup>	1.18	31%	33%	23%	14%
Web-based Online Studies (525) <sup>131</sup>	1.10	39%	25%	22%	13%

<sup>124</sup> Not all researchers code CRT responses the same way. Frederick did not indicate whether his totals included subjects failing to answer a question. Acknowledging it could inflate mean CRT scores, others exclude answers for judges failing to answer all three items. Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 14–15 n.81; Andrew J. Wistrich & Jeffrey J. Rachlinski, *How Lawyers' Intuitions Prolong Litigation*, 86 S. CAL. L. REV. 571, 586 (2013). To permit comparison with judges, we followed Guthrie et al.'s coding conventions.

<sup>125</sup> Eleven arbitrators opted not to complete CRT questions (*n* = 251; SD = 1.07). Table 1 excludes subjects failing to answer all three questions. When including non-answers, CRT score was slightly lower (M = 1.44; SD = 1.07; *n* = 251), supporting the theory that coding affects CRT scores. For the expanded sample, 25.1% got zero correct (*n* = 63), 25.5% got one correct (*n* = 64), 29.9% got two correct (*n* = 75), and 19.5% got all three items correct (*n* = 49).

<sup>126</sup> See Frederick, *supra* note 119, at 29 (reporting results from MIT and CMU).

<sup>127</sup> See Wistrich & Rachlinski, *supra* note 124, at 585–87 (evaluating lawyers from Oregon, Texas, and Ontario in the insurance sector).

<sup>128</sup> Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1499–500.

<sup>129</sup> Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 14–15.

<sup>130</sup> Frederick, *supra* note 119, at 29.

<sup>131</sup> *Id.*



While it appears that international arbitrators slightly outperformed U.S. judges on the CRT, we found a statistically meaningful difference between the arbitrators and the Florida state court judges only,<sup>132</sup> not between the arbitrators and ALJs.<sup>133</sup> These mixed results warrant caution in drawing any inferences that arbitrator reasoning is superior to judicial reasoning.<sup>134</sup> Moreover, the practical significance of any difference was minimal, as effects were statistically small.<sup>135</sup>

In addition, international arbitrators and judges were similarly likely to select the intuitive, but incorrect, responses to the CRT questions. Florida state judges nearly always selected the intuitively incorrect answer for the bat-and-ball question, and more than half of their responses for other questions were the intuitive responses suggested by the problem. Likewise, ALJs, as specialist adjudicators, tended to provide intuitive but incorrect responses.<sup>136</sup> Arbitrators' responses mimicked this pattern. Of the 158 arbitrators providing the incorrect answer on the bat-and-ball problem, 87.3% provided the intuitive answer;<sup>137</sup> on the widget problem, 62% of the 121 arbitrators providing an incorrect response identified the intuitive answer; and on the lily-pad problem, 69.8% of the 86 arbitrators providing incorrect answers gave the intuitive response as shown in Table 2.

---

<sup>132</sup> Using a test comparing correct CRT responses for 239 arbitrators and 252 Florida judges, there was a meaningful difference; and arbitrators obtained a higher a proportion of correct responses ( $\chi^2(3) = 7.92$ ;  $p = 0.048$ ;  $r = 0.13$ ;  $n = 491$ ).

<sup>133</sup> A test comparing correct total number of CRT responses for 239 arbitrators and 126 ALJs was unable to detect reliable difference ( $\chi^2(3) = 4.42$ ;  $p = 0.22$ ;  $r = 0.11$ ;  $n = 365$ ). Given the smaller ALJ sample, the null result may derive from low power. The comparison between arbitrators and judges had less than 50% power, which is below the accepted 80% threshold. Given the small effect size, sample of 781 arbitrators should have requisite power.

<sup>134</sup> Although the CRT items judges and arbitrators received were textually identical, temporal differences in administration and other factors limit the strength of inferences directly comparing judges and arbitrators. See, e.g., Maggie E. Toplak, Richard F. West & Keith E. Stanovich, *Assessing Miserly Information Processing: An Expansion of the Cognitive Reflection Test*, 20, THINKING & REASONING 147, 149 (2014) (expressing concern about use of the CRT given its increasing exposure).

<sup>135</sup> According to Cohen, effect sizes ( $r$ -values) up to 0.10 are "small," 0.11 to 0.30 are "medium," and 0.31 to 0.50 are "large." JACOB COHEN, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES 79–80 (2d ed. 1988). The effect sizes, when comparing arbitrators to U.S. judges, were close to  $r = 0.10$ . See *supra* notes 132–33.

<sup>136</sup> See Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1499–500; see also Wistrich & Rachlinski, *supra* note 124, at 587 ("Among the lawyers who got the questions wrong, 94.9 percent (149 out of 157), 58.1 percent (seventy-two out of 124), and 62.6 percent (sixty-two out of ninety-nine) chose the intuitive responses (ten cents, one hundred minutes, and twenty-four days) to the three questions, respectively.").

<sup>137</sup> For incorrect non-intuitive responses, most answers included numerical figures. Some responses, however, included written comments such as "No way to know."

Table 2: CRT Results—Subjects Providing Incorrect Responses from Samples of 239 International Arbitrators, 252 Generalist Judges, and 126 ALJs<sup>138</sup>

	Number of Participants Answering Incorrectly	Incorrect Answer Giving Intuitive Response ( <i>n</i> )	Incorrect Answer Giving Any Other Response ( <i>n</i> )
<u>Question 1:</u>		<u>10 Cents</u>	
International Arbitrators	158	87.3% (138)	12.7% (20)
Generalist Judges	181	96.7% (175) <sup>139</sup>	3.3% (6)
ALJs	79	93.7% (74)	6.3% (5)
<u>Question 2:</u>		<u>100 Minutes</u>	
International Arbitrators	121	62.0% (75)	38% (46)
Generalist Judges	141	57.5% (81)	42.5% (60)
ALJs	74	52.7% (39)	47.3% (35)
<u>Question 3:</u>		<u>24 Days</u>	
International Arbitrators	86	69.8% (60)	30.2% (26)
Generalist Judges	125	68.0% (85)	32.0% (40)
ALJs	57	63.2% <sup>140</sup> (36)	36.8% (21)

### 3. Synthesis

International arbitrators provided predominantly intuitive responses on the CRT. These results cast doubt upon narratives that arbitrators always analyze problems in fully rational ways. Like the judges who have been studied, arbitrators, as a whole, did not perform well on this relatively easy test. Although some individual arbitrators and judges showed an ability to overcome intuition with deliberation in some circumstances, members of both groups gravitated toward intuitive and impressionistic decisionmaking.

<sup>138</sup> Judge data derived from Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 15–16, and data from ALJs derived from Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1499–500, and the original dataset.

<sup>139</sup> *Blinking* incorrectly calculated the percentage as “88.4%,” but the stated proportions (“175 of 181 judges”) were accurate. Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 15–16.

<sup>140</sup> Upon reviewing the original dataset, the 64.9% reported in Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1500, was incorrect.

The CRT is not a test of legal reasoning, but scholars have identified reliable links between CRT and legal decisions.<sup>141</sup> For instance, judges performing well on the CRT performed well on an evidentiary inference problem based on *Byrne v. Boadle*.<sup>142</sup> Below, we explore whether arbitrators, like judges, tended to make intuitive judgments when confronted not with a general test like the CRT but when confronted with hypothetical disputes similar to those they confront in their professional capacity.

### B. Anchoring: Irrelevant and Relevant Anchors

Anchoring is a form of intuitive decisionmaking involving numerical estimates. When people make estimates, they tend to rely upon an initial value that is readily available, which “anchors” subsequent numerical estimations, even when the initial figure is irrelevant or an intentional distractor.<sup>143</sup> While people can adjust away from initial anchors with deliberation, they often fail to adjust sufficiently. Thus, anchors, including both reasonable and completely unreasonable anchors, often have an outsized impact on final estimates.

Kahneman & Tversky’s “wheel of fortune” experiment demonstrated the impact of irrelevant anchors on estimates.<sup>144</sup> In this classic study, researchers spun a wheel of fortune to generate a random number and then asked subjects to estimate the percentage of African states in the United Nations. Subjects’ responses were biased towards the initial value provided by the wheel of fortune, even though that number had absolutely nothing to do with African representation in the United Nations. Even when individuals are paid for assessments<sup>145</sup> and when information is updated,<sup>146</sup> anchoring persists.

---

<sup>141</sup> Jeffrey J. Rachlinski, *Processing Pleadings and the Psychology of Prejudgment*, 60 DEPAUL L. REV. 413, 420 (2011). Judges performing on the CRT did well on an evidential inference problem based on *Byrne v. Boadle*. *Id.*; see also Toplak, West & Stanovich, *supra* note 134, at 149 (“Shockingly, since it is based on just three items, the CRT has proven to be a potent predictor of performance on rational thinking tasks.”).

<sup>142</sup> We also used this hypothetical. See *infra* notes 217–24.

<sup>143</sup> See generally JENNIFER K. ROBBENOLT & JEAN R. STERNLIGHT, *PSYCHOLOGY FOR LAWYERS: UNDERSTANDING THE HUMAN FACTORS IN NEGOTIATION, LITIGATION, AND DECISION MAKING* 71–72 (2012); Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCI. 1124, 1128 (1974) [hereinafter Tversky & Kahneman, *Judgment*]; see also Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 19–21; Guthrie, Rachlinski & Wistrich, *supra* note 10, at 790–94; Wistrich, Guthrie & Rachlinski, *Disregarding*, *supra* note 77, at 1286–93.

<sup>144</sup> Tversky & Kahneman, *Judgment*, *supra* note 143.

<sup>145</sup> Gretchen B. Chapman & Eric J. Johnson, *Incorporating the Irrelevant: Anchors in Judgments of Belief and Value*, in *HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT*, 120, 125–26 (Thomas Gilovich, Dale Griffin & Daniel Kahneman eds., 2002).

<sup>146</sup> Fritz Strack & Thomas Mussweiler, *Heuristic Strategies for Estimation Under Uncertainty: The Enigmatic Case of Anchoring*, in *FOUNDATIONS OF SOCIAL COGNITION* 79, 80 (Galen V. Bodenhausen & Alan

Previous research has demonstrated that anchors affect both generalist and specialist U.S. adjudicators. One experiment asked judges to analyze damages after learning about a plaintiff who suffered serious injuries from a car crash (including several months of hospitalization and being confined later to a wheelchair) due to a truck driver's negligence. It then asked judges to rule on a motion to dismiss and assess damages. The control group was given a basic request for a damage assessment, but judges in the experimental condition were also told the defendant claimed the US\$75,000 amount-in-controversy requirement was not satisfied.<sup>147</sup> The natural anchor—US\$75,000—affected judges' damage assessments, with judges in the experimental condition awarding roughly 30% less than judges in the control condition.<sup>148</sup> In another study, ALJs assessed damages in employment discrimination where the applicable law permitted compensation for mental anguish and emotional distress. The hypothetical included testimony from the plaintiff that she suffered from "anxiety, sleeplessness, and bad dreams" and she mentioned, as an aside, that she had recently seen a "court TV show featuring a case she claimed was similar to hers."<sup>149</sup> Whereas some judges simply learned the plaintiff discussed the irrelevant TV show, others learned the compensatory damage in the "court TV show" was US\$415,300. Once again, data revealed anchors affected damage assessments; the mean award was twice as large for ALJs exposed to an irrelevant anchor.<sup>150</sup> Recent research has demonstrated irrelevant anchors likewise influenced judges from Canada, the Netherlands, and Germany.<sup>151</sup>

In transnational adjudication, relevant anchors are useful when they are grounded in fact or law, but irrelevant anchors create risk of error and injustice. In cases where the financial consequences are meaningful for one or both parties—such as damage assessments in large transnational commercial or investment disputes—irrelevant anchors pose particularly pernicious risks. We set out to examine these risks by developing two vignettes to test the impact of anchors on arbitrator decisionmaking. Whereas the first hypothetical tests

---

J. Lambert eds., 2003); Chris Guthrie & Jeffrey J. Rachlinski, *Insurers, Illusions of Judgment & Litigation*, 59 VAND. L. REV. 2017, 2026 (2006); Dan Orr & Chris Guthrie, *Anchoring, Information, Expertise, and Negotiation: New Insights from Meta-Analysis*, 21 OHIO ST. J. DISP. RESOL. 597, 597–98 (2006).

<sup>147</sup> Guthrie, Rachlinski & Wistrich, *supra* note 10, at 789–92.

<sup>148</sup> Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 21.

<sup>149</sup> Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1502–03.

<sup>150</sup> *Id.* at 1504–06.

<sup>151</sup> See Rachlinski, Wistrich & Guthrie, *Distorted Damages*, *supra* note 78; see also *id.* at 710 & n.99 (describing unpublished research demonstrating anchors influenced Taiwanese judges).

relevant anchors, the second hypothetical tests irrelevant anchors. In both scenarios, anchoring had a significant effect on outcomes.

### 1. *Relevant Anchor: Materials and Results*

One hypothetical tested two relevant anchors and also offered us an opportunity to test experimentally whether arbitrators “split the baby” when presented with two damage assessments. Nevertheless, quantitative analyses of real awards revealed that arbitrators did not render compromise awards,<sup>152</sup> but the longstanding myth persists<sup>153</sup> that arbitrators ignore the merits and “split the baby.”<sup>154</sup>

To test for the impact of a relevant anchor on arbitral decisionmaking, we created a hypothetical involving the commercialization of beachfront property where a developer created a project for low-density, ecologically responsible land development.<sup>155</sup> Later, however, the state where the land was located passed legislation prohibiting the developer from enjoying the beneficial use of the property. In the scenario, all parties—both the investor and the state—agreed a compensable expropriation occurred. The only question was damages. The vignette provided the applicable law, namely that compensation due was for the

---

<sup>152</sup> See Christopher R. Drahozal, *Busting Arbitration Myths*, 56 U. KAN. L. REV. 663, 665, 673–77 (2008) (identifying the “split the baby” myth of arbitration but providing contradictory empirical evidence); Stephanie E. Keer & Richard W. Naimark, *Arbitrators Do Not “Split the Baby”—Empirical Evidence from International Business Arbitrations*, 18 J. INT’L ARB. 573, 574–75 (2001) (analyzing arbitration awards to observe, overall, tribunals awarded roughly 47%–50% of claimed amounts but identifying that the average figure masked a bimodal distribution where tribunals either rendered awards in favor of either claimant or respondent); Carter Greenbaum, *Putting the Baby to Rest: Dispelling a Common Arbitration Myth*, 26 AM. REV. INT’L ARB. 101, 101 (2015) (providing empirical data that “the incidence of compromise awards in commercial arbitration is insignificant”).

<sup>153</sup> See *supra* note 7; DOUGLAS SHONTZ, FRED KIPPERMAN & VANESSA SOMA, RAND INST. FOR CIV. JUST., BUSINESS-TO-BUSINESS ARBITRATION IN THE UNITED STATES: PERCEPTIONS OF CORPORATE COUNSEL, at x, 7–12 (2011), [http://www.rand.org/content/dam/rand/pubs/technical\\_reports/2011/RAND\\_TR781.pdf](http://www.rand.org/content/dam/rand/pubs/technical_reports/2011/RAND_TR781.pdf) (identifying that parties’ “overwhelming believe that arbitrators tend to ‘split the baby’ with their rulings—that is, they are unwilling to rule strongly for one party”). The history, scope, strength, and persistence of the “split the baby” myth is a topic of quantitative analysis beyond the scope of this Article. We nevertheless observe that many commentators continue to discuss this problem. See, e.g., Zela G. Claiborne, *Top Five Myths about Commercial Arbitration*, JAMS (Sept. 7, 2015), <https://www.jamsadr.com/publications/2015/top-five-myths-about-commercial-arbitration>; AM. ARB. ASS’N, SPLITTING THE BABY: A NEW AAA STUDY (2007), [https://www.adr.org/aaa/ShowPDF?doc=ADRSTG\\_014040](https://www.adr.org/aaa/ShowPDF?doc=ADRSTG_014040); Ana Carolina Weber et al., *Challenging the “Splitting the Baby” Myth in International Arbitration*, 31 J. INT’L ARB. 719 (2014).

<sup>154</sup> Drahozal, *supra* note 152, at 675; Chris Guthrie, *Misjudging*, 7 NEV. L.J. 420, 454 (2007).

<sup>155</sup> To enhance the external validity of the scenario, we patterned it after other cases confronting relevant anchors to assess the value of beach front property. *Compañía del Desarrollo de Santa Elena, S.A. v. Costa Rica*, ICSID Case No. ARB/96/1, Final Award (Feb. 17, 2000); *Unglaube v. Costa Rica*, ICSID Case Nos. ARB/08/1, ARB/09/20, Award (May 16, 2012).

“fair market value”, or “what a willing buyer would pay to a willing seller for the best use of the property.” Arbitrators were then instructed to resolve the dispute as if they were a sole arbitrator. We told arbitrators that both parties submitted credible expert reports referencing historical data from real estate sales, listing prices of the original property, sunk costs, market data on the demand for beach rental property, and the potential profits of ecologically-friendly developments in similar states. The only difference between the scenarios subjects received was the valuation in expert reports. In all versions, the respondent state that had expropriated the property asserted the land value was US\$1 million. In the low anchor condition, the developer claimed damages of US\$10 million; and in the high anchor condition, the developer claimed US\$50 million.

Measures of central tendency reflected the influence of relevant anchors. While the mean award for all answers was roughly US\$16 million,<sup>156</sup> there was meaningful variation in damage assessments for low and high anchors.<sup>157</sup> Arbitrators in the high anchor condition made reliably larger awards. Whereas the average award in the high anchor condition was nearly US\$24.8 million,<sup>158</sup> the mean award for arbitrators in the low anchor condition was roughly US\$5.8 million.<sup>159</sup>

At first blush, damage awards might appear to support a “split the baby” hypothesis, as—for both conditions—mean awards were roughly halfway between the values of the two expert reports. The difference between the competing expert reports in the low anchor condition should generate a compromise award of US\$5.5 million,<sup>160</sup> which was similar to the US\$5.8 million mean. Likewise, splitting the difference between the experts in the high

---

<sup>156</sup>  $M = 16,430,556$ ;  $n = 90$ ;  $SD = 16,942,692$ .

<sup>157</sup> A  $t$ -test revealed meaningful variation ( $t(96) = -6.844$ ;  $p < 0.001$ ;  $r = 0.57$ ;  $n = 98$ ). It was not necessary to transform damages, as skewness (1.06) was acceptable. Results remained significant using a non-parametric Mann-Whitney U-test of medians ( $U = 1548$ ;  $p < 0.001$ ). The smaller  $n$  reflects subjects randomly received either a beach-front anchoring hypothetical or settlement framing hypothetical.

<sup>158</sup>  $M = 24,773,585$ ;  $n = 53$ ;  $SD = 18,314,970$ . For the high anchor, the 25th percentile was US\$ 7,500,000 the median was US\$25,000,000, and the 75th percentile was US\$44,500,000.

<sup>159</sup>  $M = 5,794,444$ ;  $n = 45$ ;  $SD = 3,451,486$ . For the low anchor, the 25th percentile was US\$2,500,000, the median was US\$50,000,000, and the 75th percentile was US\$10,000,000.

<sup>160</sup> The difference between the two expert reports was US\$9 million, so 50% is US\$4.5 million. By adding US\$4.5 million to the state's US\$1 million valuation or subtracting US\$4.5 million from the developer's US\$10 million claim creates a compromise award of US\$5.5 million.

anchor condition would create a US\$25.5 million<sup>161</sup> award, which was similar to the reported US\$24.8 million mean. Further analysis, however, reveals the visual similarity is an oversimplification that hides fundamental variance in arbitrator decisionmaking.

We used arbitrators' responses to calculate whether an award was a compromise between the two expert valuations. This meant an arbitrator rendered an award precisely 50% between the two expert valuations, when it rendered a US\$5.5 million award in the low anchor condition or US\$25.5 million in the high anchor condition. Where arbitrators awarded US\$1 million, this was a 0% award, as damages fully reflected respondent's expert. When damages were in line with claimant's expert, the arbitrator awarded 100%. Other proportions reflected arbitrators' awards and degree of compromise.<sup>162</sup>

Irrespective of experimental condition, tests were unable to identify meaningful differences in subjects' proportionate responses.<sup>163</sup> In other words, there was no evidence that anchoring affected the relative proportions awarded. This, in turn, suggests arbitrators' propensity to award compromise awards was equivalent across conditions.<sup>164</sup>

The responses revealed variation in damage assessments. Figure 1 reflects less than half of arbitrators (41.8%;  $n = 41$ ) awarded investors *more than 50%* of the claimed damages; and as reflected by red shades, more than half arbitrators (56.1%;  $n = 55$ ) awarded investors *less than 50%* of claimed damages. Only 2% ( $n = 2$ ) of arbitrators rendered pure compromise award that precisely "split the baby" between the two expert reports. Instead, one of the largest sub-groups of responses came from arbitrators awarding 40%–49% of damages.<sup>165</sup>

---

<sup>161</sup> The difference between the two reports was US\$49 million, so 50% is US\$24.5 million. Adding US\$24.5 million to the state's US\$1 million valuation or subtracting US\$24.5 million from the developer's US\$50 million claim creates a compromise award of US\$25.5 million.

<sup>162</sup> To calculate the percentage, we subtracted US\$1 million from awarded damages (to address respondent's concession of a US\$1 million valuation). For the low anchor, we divided that amount by US\$9 million; for the high anchor, we divided by US\$49 million, the respective spreads between the two reports.

<sup>163</sup> The proportions exhibited acceptable skewness (0.03) and required no transformation. With two experimental conditions, a *t*-test analyzed group differences in the high and low anchor groups; and the test was unable to identify a meaningful difference ( $t(96) = 0.625$ ;  $p = 0.53$ ;  $r = 0.06$ ;  $n = 98$ ).

<sup>164</sup> The analysis lacked sufficient power to conclude there was no effect of anchoring on proportion awarded. Given the small effect size ( $r = 0.06$ ), *a priori* power analysis reveals a sample of over 781 arbitrators would be required to make inferences about a null result.

<sup>165</sup> Five arbitrators awarded 49%, two arbitrators awarded 48%, and ten arbitrators awarded 44%. Five arbitrators awarded 39%. One awarded 33%.

The two largest responses did not reflect a compromise between the parties' claims. Rather, the most common response involved arbitrators taking an “all-or-nothing” approach and following the expert report of one party. Leaving aside the 5% of arbitrators who gave less than either party anticipated, 13% ( $n = 13$ ) of arbitrators rendered awards following respondent's expert report, and 26.5% ( $n = 26$ ) of arbitrators used claimant's expert report as the basis of damages. These responses undermine claims that arbitrators, as a group, “split the baby.”

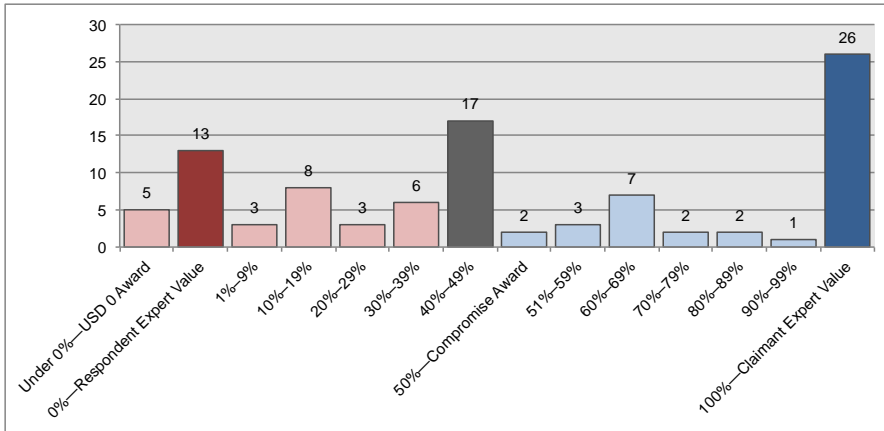


Figure 1. Response Frequency: Arbitrators' Proportionate Damage Awards ( $n = 98$ ).

While the data did not demonstrate arbitrators never “split the baby”—arbitrators sometimes did precisely that—it does cast doubt on the universality and prevalence of such a narrative. It is empirically wrong to suggest that international arbitrators, as a group, exhibited monolithic tendencies. Rather, the data suggest that conventional wisdom about arbitrator decisionmaking must be replaced with an alternative theory.<sup>166</sup> At best, there were three different paradigms for international arbitrators, namely a group inclined to prioritize claimant valuations, a group inclined to prioritize respondent valuations, and a

<sup>166</sup> Judge Posner argues arbitrators seek to maximize appointments by rendering compromise awards. Posner, *supra* note 7, at 1260; Richard Posner, *What Do Arbitrators Maximize?*, in *LAW AND ECONOMICS OF INTERNATIONAL ARBITRATION: FIFTH INTERNATIONAL CONFERENCE ON LAW AND ECONOMICS AT THE UNIVERSITY OF ST. GALLEN* 123, 124–25 (Peter Nobel & Philipp von Ins eds., 2014) (“[T]here would be a tendency of arbitrators to split the difference between the parties rather than side entirely with one party.”); see also Robert D. Cooter, *The Objectives of Private and Public Judges*, 41 *PUB. CHOICE* 107, 110, 128 (1983) (exploring the theoretical rational actor model). The data disrupted this theory, as a small group of arbitrators rendered “split the baby” awards. More arbitrators rendered “all or nothing” or somewhat more respondent-favorable awards, suggesting an alternative theory is warranted to explain intuitive adjudication styles.



group that was roughly in the middle but tended to render more respondent-favorable awards. It is, however, difficult to predict such propensities in advance or how they might vary by context. Our hypothetical involved an ITA dispute; it is possible—but by no means certain—that arbitrators behave differently in ICA disputes.<sup>167</sup> At a minimum, given the control parties have over the appointment of one arbitrator, the data suggest parties should engage in due diligence when making arbitrator appointments. Parties may even wish to consider providing arbitrators in advance with cognitive assessments to identify their intuitive predispositions and the capacity to override intuitive reasoning with rationality.

## 2. *Irrelevant Anchor: Materials and Results*

One would hope that relevant anchors that have a nexus with the applicable law and facts influence damage awards, as anchoring has the capacity to be adaptive. The data demonstrated that the relevant anchors exerted this influence, and none of the arbitrators made a damage award in excess of an investors' expert report.<sup>168</sup> This finding left open the question, however, of whether irrelevant anchors—which risk improperly inflating damages—also influence decisionmaking.

To test the influence of *irrelevant* anchors, we created a hypothetical dispute involving a small transnational law firm (Law Firm) that, at its clients' behest, opened a local office in a new country. The host country, however, strictly regulated the practice of law by foreign lawyers. A government raid on Law Firm resulted in destruction of irreplaceable items (including a rare family photo album and specialized technological equipment), incarceration of employees, and deportation of foreign lawyers.<sup>169</sup> Law Firm experienced adverse publicity, damaged reputation, and lost clients. Law Firm then initiated ITA to recover damages, claiming the host state failed to provide "fair and equitable treatment"

---

<sup>167</sup> Studies regarding ICA, whereby commercial arbitrators also did not demonstrate a pure propensity to "split the baby" in real cases, cast a degree of doubt on such a hypothesis. *Compare supra* notes 152–53, with Figure 1.

<sup>168</sup> We note, however, that five of the arbitrators awarded an amount for expropriation less than what the respondent state conceded was due. *See* Figure 1. It is possible that these arbitrators had an intuitive approach favoring states, did not closely read the question, or there was some other basis for the assessment.

<sup>169</sup> Aspects of the hypothetical were similar to other disputes resolved by arbitration. *Al-Kharafi & Sons Co. v. Libya*, (Kuwait v. Libya), Final Arbitral Award, 4–5 (Mar. 22, 2013); *Desert Line Projects LLC v. Yemen* (Oman v. Yemen), ICSID Case No. ARB/05/17, Award, 4–10 (Feb. 6, 2008); *Mitchell v. Democratic Republic of the Congo* (U.S. v. Democratic Republic of the Congo), ICSID Case No. ARB/99/7, Decision on the Application for Annulment of the Award, 3 (Nov. 1, 2006).

(FET). The Request for Arbitration requested moral damages, namely damages for inchoate harms like duress, loss of reputation, or roughly equivalent to pain and suffering.<sup>170</sup> We asked arbitrators to assume they were the tribunal chair, there was jurisdiction over the dispute, there was an FET violation, and the calculation of damages rested on equitable principles under the applicable international law standards.

Unbeknownst to arbitrators, they did not all receive the same version. Like the experiment including information about an irrelevant “court TV show,” we injected information about an irrelevant case<sup>171</sup> with no bearing on the dispute. Reflecting a common practice in international arbitration, the hypothetical indicated: “After the final hearing, you go to dinner with your co-arbitrators to relax. During dinner, as an aside, one of your co-arbitrators mentions a case where a U.S. District Court applied U.S. domestic tort law to hold Sudan liable for \_\_\_\_ to those injured by the bombing of a ship in Yemen.” We randomly assigned arbitrators to one of four conditions. In the control group, arbitrators received no information about the valuation of the unrelated U.S. tort case. In contrast, three experimental conditions contained different valuations of the irrelevant case, namely: (a) US\$1 million in damages (a low anchor), (b) US\$50 million in damages (a medium anchor), or (c) US\$300 million in damages (a high anchor). Arbitrators were asked: “How much do you award Law Firm in moral damages for the FET violation?”

Irrelevant anchors influenced international arbitrators’ damage awards. Using raw data, the average award was US\$9.2 million,<sup>172</sup> but damages varied across conditions. The mean award in the control condition was roughly US\$10 million.<sup>173</sup> Mean awards in the low anchor condition were roughly US\$5 million,<sup>174</sup> and US\$5.5 million in the medium anchor condition.<sup>175</sup> The mean raw award in the high anchor condition was US\$16.3 million.<sup>176</sup> Median awards likewise reflected the influence of anchors. For both the control group and low-

---

<sup>170</sup> See Matthew T. Parish, Annalise K. Nelson & Charles B. Rosenberg, *Awarding Moral Damages to Respondent States in Investment Arbitration*, 29 BERKELEY J. INT’L L. 225, 225–30 (2011); Ben Saul, *Compensation for Unlawful Death in International Law: A Focus on the Inter-American Court of Human Rights*, 19 AM. U. INT’L L. REV. 523, 555–60 (2004).

<sup>171</sup> The irrelevant anchor was based upon a case involving U.S. sailors injured during a bombing. *Harrison v. Republic of Sudan*, 882 F. Supp. 2d 23 (D.D.C. 2012).

<sup>172</sup>  $M = 9,168,2485; n = 218; SD = 29,366,890.$

<sup>173</sup>  $M = 10,347,348; n = 49; SD = 36,177,797.$

<sup>174</sup>  $M = 4,975,636; n = 55; SD = 18,903,177.$

<sup>175</sup>  $M = 5,478,068; n = 59; SD = 10,744,534.$

<sup>176</sup>  $M = 16,269,091; n = 55; SD = 41,659,270.$

anchor group, Figure 2 demonstrates the median award was US\$100,000. By contrast, the median award for both the medium anchor and high anchor groups was US\$1,000,000.

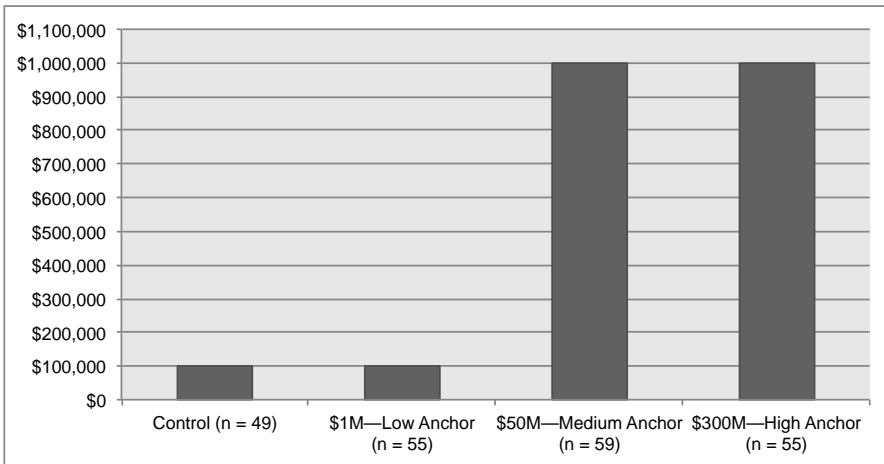


Figure 2. Median Award: Legal Services Hypothetical with Four Anchors.

In an effort to identify the nuance in the variation of damage awards, Table 3 offers a quartile distribution of the different awards across conditions. Particularly in the 75th percentile, medium and high anchors appeared to swing damages in an upward direction.

Table 3: Quartile Distributions of Damage Awards in Anchoring Conditions.

	25th Percentile	Median—50th Percentile	75th Percentile	Total Sample
<u>Control Condition:</u> No Anchor	0	100,000	1,000,000	49
<u>Low Anchor Condition:</u> US\$1 million	0	100,000	1,000,000	55
<u>Medium Anchor Condition:</u> US\$50 million	1,000	1,000,000	5,000,000	59
<u>High Anchor Condition:</u> US\$300 million	0	1,000,000	10,000,000	55

To identify whether variations in damages were meaningful, we conducted a between-groups Analysis of Variance (ANOVA).<sup>177</sup> Because of positive skewing, we transformed the data<sup>178</sup> to ensure it met the assumptions of statistical tests.<sup>179</sup> The test revealed anchoring exerted a statistically significant effect on international arbitrators' damage awards.<sup>180</sup>

Not all anchors had the same effect, however. Conservative follow-up comparisons<sup>181</sup> identified that the high anchor was reliably influential. Damage awards in the high anchor condition were meaningfully larger than awards from arbitrators in either the control group or the low anchor condition.<sup>182</sup> A more liberal test<sup>183</sup> suggested both the high and medium anchor were linked with increased damages.<sup>184</sup> It was never possible, however, to detect a reliable difference between awards in the control group and low anchor condition.<sup>185</sup> The

---

<sup>177</sup> See TIMOTHY C. URDAN, *STATISTICS IN PLAIN ENGLISH 105–10* (3d ed. 2010) (explaining ANOVAs and their proper use).

<sup>178</sup> Winsorizing requires identifying and converting extreme values into the upper and lower bounds of the distribution. W.J. Dixon, *Simplified Estimation from Censored Normal Samples*, 31 *ANNALS MATH. STAT.* 385, 385 (1960); John W. Tukey, *The Future of Data Analysis*, 33 *ANNALS MATH. STAT.* 1, 18–19 (1962). Winsorizing identifies outliers using Tukey's hinges, which computes low and high cutoffs, and replaces outlying values with the upper and lower bounds of Tukey's hinges. This reformulates data to fit test assumptions but retains data. DAVID J. SHESHKIN, *HANDBOOK OF PARAMETRIC AND NONPARAMETRIC STATISTICAL PROCEDURES* 403 (3d ed. 2004); Franck, *supra* note 82, at 456.

<sup>179</sup> Skewness of the raw data was an unacceptable 5.17. After Winsorization, skewness was an acceptable 0.92.

<sup>180</sup> The ANOVA results were significant ( $F(3217) = 4.696; p = 0.003; r = 0.25; n = 218$ ). A non-parametric Kruskal-Wallis test was marginally significant ( $\chi^2(3) = 7.203; p = 0.06; r = 0.18; n = 218$ ). When combining the control and the low anchor conditions, which appeared to operate similarly, a Kruskal-Wallis test revealed a significant group difference ( $\chi^2(2) = 6.755; p = 0.03; r = 0.17; n = 218$ ). When combining the medium and high anchor conditions, which appeared to operate similarly, a Kruskal-Wallis test revealed a significant group difference ( $\chi^2(2) = 7.166; p = 0.03; r = 0.18; n = 218$ ).

<sup>181</sup> Tukey's honestly significant difference (HSD) provides follow-up significance testing. FREDERICK J. GRAVETTER & LARRY B. WALLNAU, *ESSENTIALS OF STATISTICS FOR THE BEHAVIORAL SCIENCES* 365 (6th ed. 2008).

<sup>182</sup> HSD comparisons between the high anchor and: 1) the control group ( $p = 0.03$ ) or 2) the low anchor ( $p = 0.01$ ) were meaningful. We could not find a meaningful difference when comparing awards in medium and high anchor groups ( $p = 0.81$ ).

<sup>183</sup> A Fisher's Least Significant Difference (LSD) test permits comparison of sub-groups for individual group differences. LSD, however, is more likely to identify meaningful differences when compared to more conservative HSD analyses.

<sup>184</sup> For LSD comparisons using the medium anchor, the significant effect was comparing the medium anchor and low anchor ( $p = 0.02$ ). Comparing the medium anchor with the control group was marginally significant ( $p = 0.05$ ). Comparisons between medium and high anchors remained non-significant ( $p = 0.37$ ).

<sup>185</sup> For HSD comparisons between control and low anchor conditions, there was no significant effect ( $p = 0.99$ ); and for LSD, there was no identifiable effect ( $p = 0.74$ ). Because of the proportion of responses in the control condition where arbitrators rendered awards that were below the value provided in the "low" anchor condition, these results have limited value in identifying the lack of an effect of a "low" anchor. Moreover, the

significant results about high anchors nevertheless opens the possibility that, in international arbitration, anchors must be sufficiently large to exert a meaningful influence.

The results demonstrated that, like generalist and specialist judges, anchoring influences international arbitrators. It is, unfortunately, not possible to compare the influence of anchoring on arbitrators and judges directly, as hypotheticals provided to the groups involved distinct legal questions, different anchors, and different categories. Nevertheless, we can conclude with some confidence that anchoring influenced both groups.

The results suggest that it may be prudent to explore reforming international arbitration to require parties to plead damages with specificity at an early phase (or otherwise provide detailed expert reports in advance) to justify amounts claimed. As claimants are positioned to know their own damages and relevant information is within their control, the burden of such a requirement is likely minimal; and the potential benefit of creating procedural mechanisms to minimize the pernicious effect of anchoring likely outweighs systemic costs. Moreover, as ensuring there are thorough assessments of damages at the outset of a case is consistent with best practices in international arbitration, it is likely that quality counsel already conduct these analyses and consult with their clients in advance about the costs-and-benefits of pursuing arbitration. Imposing formal requirements—whether in institutional rules or party agreement—will therefore not change the practice of many lawyers and will create incentives for quality in others.

We also note that some of the international arbitrators were not fully content with the brief information we provided when responding to the anchoring questions. Among those responding to the beachfront property hypothetical testing relevant anchors, 25% complained about insufficient information; among those responding to the legal services hypothetical testing irrelevant anchors, 18.8% of arbitrators made manuscript comments complaining about insufficient information to make a calculation. For the beachfront property dispute, one subject noted that, “much depends upon quality of experts” and identified concerns about the subjective aspects of what is a “willing buyer.” This, however, also makes them somewhat similar to U.S. judges, as similar proportions of judges who received one-page anchoring hypotheticals also expressed discontent with the sufficiency of material provided.

---

lack of a statistically significant effect means that drawing reliable inferences about the absence of an effect is problematic.

Anchoring is a uniquely powerful phenomenon, but other information—including other numbers, like the amounts claimed or expert reports with relevant values—may minimize anchoring's deleterious effects. Moreover, arbitrators in live cases will have much more information about the parties and the dispute. Opposing counsel will likely use opposing anchors in the course of their advocacy, which one would hope would offer a relevant anchor based upon the fact and law. International arbitrators also have the authority to ask questions to test the integrity of anchors and to appoint experts. Even where anchors exert a powerful influence, effects could be muted using existing procedures within arbitration as debiasing tools.<sup>186</sup>

Although there are inevitable limitations, there are reasons to believe the results are generalizable beyond the laboratory. Damage demands, whether by claimants or respondents in a counter-claim, are salient anchors. Other research has demonstrated that, even with more information or alternative anchors, adjudicators were unable to fully disregard initial anchors. International arbitrators do have time and opportunity to deliberate, but it is unclear whether deliberation will ameliorate the effects of an anchor by encouraging deliberation or exacerbate the effects of an anchor through group polarization.

### C. Framing

The framing of options can influence how people make decisions.<sup>187</sup> When choosing between options that appear to be gains relative to the status quo, people become risk avoiders, preferring a sure thing to a gamble. When choosing between options that seem like losses relative to the current state of affairs, people often make risk-seeking choices, rejecting a certain loss in favor of a

---

<sup>186</sup> Normative reforms deriving from evidence-based insights are discussed in Section IV. Debiasing in anchoring is notoriously difficult, as inoculants can create alternative anchors or facilitate over-correction. See Robert A. Prentice, *Chicago Man, K-T Man, and the Future of Behavioral Law and Economics*, 56 VAND. L. REV. 1663, 1757 (2003); Rachlinski, Wistichrich & Guthrie, *Distorted Damages*, *supra* note 78, at 732–35; Jeffery J. Rachlinski, *A Positive Psychological Theory of Judging in Hindsight*, 65 U. CHI. L. REV. 571, 603 (1998).

<sup>187</sup> See, e.g., Daniel Kahneman & Amos Tversky, *Choices, Values and Frames*, 39 AM. PSYCHOLOGIST 341 (1984); Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision Under Risk*, 47 ECONOMETRICA 263 (1979); Amos Tversky & Daniel Kahneman, *The Framing of Decisions and Psychology of Choice*, 211 SCI. 453 (1981). But see James N. Druckman, *Using Credible Advice to Overcome Framing Effects*, 17 J.L. ECON. & ORG. 62 (2001) (suggesting framing can diminish or disappear when subjects obtain credible information).

gamble that might enable them to avoid losing anything at all. People find losses more aversive and unfair than they find equivalent gains attractive and fair.<sup>188</sup>

In one illustrative study, researchers gave subjects the following problem: “A company is making a small profit. It is located in a community experiencing a recession with substantial unemployment . . . .”<sup>189</sup> In the loss condition, subjects learned there was “no inflation,” but the company was decreasing wages by 7%. In the gain condition, subjects learned that there is “substantial unemployment and inflation of 12%,” but the company decided “to increase salaries by only 5%.”<sup>190</sup> Researchers asked subjects to evaluate the fairness of these options. Although employees’ real income shifts were approximately the same, “judgments of fairness are strikingly different. A wage cut is coded as a loss and consequently judged unfair. A nominal raise which does not compensate for inflation is more acceptable because it is coded as a gain to the employee . . . .”<sup>191</sup>

Researchers have found that framing influences both novice and expert decisionmakers, including judges.<sup>192</sup> We set out to examine whether it also influences international arbitrators. We used two groups of framing scenarios to test international arbitrators. The first set of hypotheticals involved more traditional gain versus lost framing effects, which used a standard methodology to request either a numerical price adjustment or a fairness adjustment. The second hypothetical involved framing deriving from a contract rescission, which also permitted exploration of the potential effect of arbitrator appointment on commercial disputes.

---

<sup>188</sup> Amos Tversky & Daniel Kahneman, *Advances in Prospect Theory: Cumulative Representation of Uncertainty*, 5 J. RISK & UNCERTAINTY 297, 307–08 (1992). Low-probability losses and gains can operate differently. See Chris Guthrie & Jeffrey J. Rachlinski, *Insurers, Illusions of Judgment & Litigation*, 59 VAND. L. REV. 2017, 2034–35 (2006).

<sup>189</sup> Daniel Kahneman, Jack L. Knetsch & Richard Thaler, *Fairness as a Constraint on Profit Seeking: Entitlements in the Market*, 76 AM. ECON. REV. 728, 731 (1986).

<sup>190</sup> *Id.*

<sup>191</sup> *Id.* at 731–32.

<sup>192</sup> See Linda Babcock et al., *Forming Beliefs About Adjudicated Outcomes: Perceptions of Risk and Reservation Values*, 15 INT’L REV. L. & ECON. 289, 293–97 (1995) (framing affects lawyers); Guthrie, Rachlinski & Wistrich, *supra* note 10, at 796–97 (framing affects judges); Barbara J. McNeil et al., *On the Elicitation of Preferences for Alternative Therapies*, 306 NEW ENG. J. MED. 1259, 1262 (1982) (framing affects physicians); Devon G. Pope & Maurice E. Schweitzer, *Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes*, 101 AM. ECON. REV. 129, 155 (2011) (framing affects professional golfers).

### 1. *Framing: Price Adjustment and Fairness Assessment*

We tested whether arbitrators evaluated gain and loss frames differently using a scenario involving an international price renegotiation dispute.<sup>193</sup> In our materials, two commercial entities—Outsourcer and Service Provider—had a contract for IT-related services. The contract required periodic price adjustment, but permitted pricing disputes to be resolved with arbitrators deciding *ex aequo et bono*, which requires decisions based upon fairness and equitable principles, and is a well-known basis for decision in international arbitration.<sup>194</sup>

We randomly assigned arbitrators to different conditions to explore the effect of framing. In the loss version, arbitrators learned, “A group of prominent economists predict that the economic outlook is muted with an inflation rate of 0%. Outsourcer proposes to cut its pay to Service Provider by 3%.” In the gain version, arbitrators learned: “A group of prominent economists predict that the economic outlook is bustling with an inflation rate of 5%. Outsourcer proposes to increase its pay to Service Provider by 2%.” In both scenarios, the net economic impact to Service Provider is a 3% difference in revenues.

Having presented the arbitrators with one of these two basic substantive frames, we asked two different questions requiring them to assess the dispute.<sup>195</sup> First, we asked a randomly assigned group of arbitrators to make a price adjustment. We asked a second randomly assigned group of arbitrators to evaluate the relative fairness of Outsourcer’s proposed price adjustment on a four-point scale (from “completely fair” to “very unfair”).

We found that gain and loss frames affected arbitrators’ evaluations.

When asked to make price adjustments, the arbitrators in the gain condition adjusted, on average, 4.44%, while those in the loss condition adjusted by an

---

<sup>193</sup> Price review disputes, for example, are typical within the oil and gas industry. *See, e.g.*, Julian Cardenas Garcia, *An Era of Petroleum Arbitration Mega Cases*, 35 HOUS. J. INT’L L. 537, 539 (2013); Christopher Goncalves, *Breaking Rules and Changing the Game: Will Shale Gas Rock the World?*, 35 ENERGY L.J. 225, 251 (2014); *Gas Price Renegotiation: A Sign of the Times*, WINSTON & STRAWN LLP (Jan. 21, 2015), <http://cdn2.winston.com/images/content/9/2/v2/92799/Gas-Price-Renegotiation-JAN2015.pdf>; *see also* Suez v. Argentina, ICSID Case No. ARB/03/19, Decision on Liability, ¶ 83 (July 30, 2010), <http://www.italaw.com/sites/default/files/case-documents/ita0826.pdf> (discussing price review disputes within the water distribution and waste water treatment context).

<sup>194</sup> Statute of the International Court of Justice art. 38(2), June 26, 1945, 59 Stat. 1055, 1060, 3 Bevans 1153, 1187; Trakman, *supra* note 35, at 631–32.

<sup>195</sup> One subset of arbitrators randomly received the price adjustment version and was randomly assigned to the gain or loss condition; and another subset of arbitrators randomly received the fairness assessment version and was randomly assigned to the gain or loss condition.



average of only 1.03%<sup>196</sup> as shown in Table 4. Stated simply, arbitrators in the gain condition adjusted prices four times as much as those in the loss condition.

Table 4: Percentage of Price Adjustment when Framed as Loss or Gain.

	Mean	Standard Deviation	Total
<u>Gain Condition:</u> 5% Inflation and 2% Raise	4.444	1.838	65
<u>Loss Condition:</u> 0% Inflation and 3% Cut	1.030	3.414	50

Likewise, we also used the same Outsourcer and Service Provider scenario to request fairness assessments. Rather than asking for a numerical decision, as international arbitration law sometimes requires arbitrator to make assessments based upon equitable principles, we asked the arbitrators to assess the fairness of the modification. The results, illustrated in Table 5, demonstrated that framing affected arbitrators' fairness assessments.

Table 5: Fairness Evaluation of Price Disputes when Framed as Loss or Gain: Percentage Giving Classification ( $n$  = number of responses).

	Completely Fair	Acceptable	Unfair	Very Unfair	Total
<u>Gain Condition:</u> 5% Inflation and 2% Raise	6.3% (4)	39.7% (25)	41.3% (26)	12.7% (8)	63
<u>Loss Condition:</u> 0% Inflation and 3% Cut	1.6% (1)	25.4% (16)	65% (41)	8% (5)	63

Because so few arbitrators in both conditions identified Outsourcer's proposal as "Completely Fair," we condensed the four-category variable into a two-category variable. We combined responses evaluating the proposal as fair—namely, Completely Fair and Acceptable—and responses evaluating the

<sup>196</sup> A  $t$ -test analyzed identified reliable difference in price adjustments ( $t(135) = -6.875; p < 0.001; r = 0.54; n = 115$ ). Using Cohen's conventions, the effect size was small-to-medium. See generally COHEN, *supra* note 135, 113–16.

proposal as unfair—namely Unfair and Very Unfair, and then tested for differences. We found evidence that framing affected arbitrators' decisions.<sup>197</sup>

In short, arbitrators, like judges, were susceptible to the effects of framing. Direct comparisons are not available, but based on a review of the prior studies on judges,<sup>198</sup> it appears that arbitrators were not more susceptible to framing.

## 2. Framing: Contract Rescission and Appointment

Framing creates asymmetries between parties in transactions (e.g., buyers versus sellers), as well as parties embroiled in disputes (e.g., claimants versus respondents). These natural frames might influence judges and arbitrators.<sup>199</sup>

In an unpublished study, Rachlinski and Wistrich found that framing affected judges' decisions.<sup>200</sup> They gave Utah state judges a vignette involving a dispute over a video game sale. Both litigants, who were attending a video game convention, misunderstood the game's value when concluding the contract. The only issue was whether there was a mutual mistake about a basic assumption of the contract warranting rescission. The hypothetical was analogous to *Sherwood v. Walker*, where a mutual mistake by both parties voided a contract.<sup>201</sup> In *Sherwood*, the issue was whether the cow was fertile or infertile; and in the hypothetical, the question was whether the goods being sold was a rare vintage game or not.<sup>202</sup>

---

<sup>197</sup> Condensing the categories into a 2x2 design, a Fisher's exact test revealed that gain and loss frames reliably influenced arbitrators' fairness assessments ( $p = 0.04$ ;  $r = 0.20$ ;  $n = 126$ ). Using Cohen's conventions, the effect size was small-to-medium. *See id.*

<sup>198</sup> *See* Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1507–09; Rachlinski, Guthrie & Wistrich, *supra* note 105, at 1240–41. For example, ALJs assessed framing in a different context using identical categories. Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1507–09. When assessing economically equivalent rent payments framed as a gain (i.e., a discount) or a loss (i.e., a surcharge), framing exerted a reliable effect on ALJs. *Id.* For judges in the gain condition responding to rent payments, 29% evaluated the payment as “Completely Fair,” 67% evaluated rent payment as “Acceptable,” 5% evaluated the assessment as “Unfair,” and 0% ranked the assessment “Very Unfair.” *Id.*; *compare id.*, with Table 4.

<sup>199</sup> *Cf.* Tess Wilkinson-Ryan & David A. Hoffman, *The Common Sense of Contract Formation*, 67 STAN. L. REV. 1269 (2015) (experimental research on ordinary individuals in the United States reflected that intuitive predispositions affected assessments of contract formation and also revealed a gap between existing U.S. contract doctrine and colloquial understandings of contracts).

<sup>200</sup> Jeffrey J. Rachlinski & Andrew J. Wistrich, *Gains, Losses, and Judges: Framing and the Judiciary* (Apr. 2017) (unpublished manuscript) (on file with authors).

<sup>201</sup> 33 N.W. 919 (Mich. 1887).

<sup>202</sup> *Id.* at 923–24. Although a classic in contracts casebooks, *Sherwood* is of limited value in Michigan given *Lenawee County Board of Health v. Messerly*, 331 N.W.2d 203 (Mich. 1981).

Not all judges received the same version of the hypothetical. In one version, both parties believed the video game had little value and the plaintiff was the *seller* who sold the video game for US\$1 but later learned it was worth US\$38,000. In the second version, both parties believed the video game had high value and the plaintiff was the *buyer* who bought the videogame for US\$38,000 but later learned it was worth US\$1.<sup>203</sup> Both scenarios asked judges to decide whether to grant plaintiff's summary judgment motion for rescission. Although the applicable law arguably required rescission in both cases,<sup>204</sup> framing influenced judges' willingness to rescind contracts. When the plaintiff was the buyer (who paid US\$38,000 for a worthless game), 82.3% (14 out of 17) of judges rescinded the contract.<sup>205</sup> In contrast, when the plaintiff was the seller (who received US\$1 for a valuable game), only 40.6% (13 out of 32) of judges rescinded the contract.<sup>206</sup>

We developed an analogous scenario for the arbitrators participating in our study.<sup>207</sup> Rather than a video game contract, the dispute involved a five-year concession contract to extract "all minerals" on a 2000 hectare site.

In one version, the concession was for a site named "LaKapa," where both parties believed the site contained iron pyrite (i.e., fool's gold) and agreed on a contract price of US\$1 million. An independent expert appointed by both parties valued the site at US\$2.5 million but mistakenly surveyed Lake Apa (i.e., the

---

<sup>203</sup> This hypothetical is a slight variation on classic gain/loss framing. Selling a valuable videogame for US\$1 is the equivalent of a foregone gain where the seller obtained some value but nevertheless did not obtain the value both parties believed to exist. By contrast, buying a videogame for US\$1 that both parties believed was worth US\$38,000 is a loss. Rachlinski & Wistrich, *supra* note 201.

<sup>204</sup> As the hypothetical invoked the *Restatement (Second) of Contracts*, it is possible that rescission would not be granted. The judges were told: "Utah courts have adopted the rule regarding mutual mistake stated in the *Restatement (Second) of Contracts*, which provides that a contract is voidable when 'a mistake of both parties at the time a contract was made as to a basic assumption on which the contract was made has a material effect on the agreed exchange of promises.'" The judges were not instructed on other *Restatement* provisions, including the full text of § 152 or § 154. Those provisions—involving risk allocation, which party bears the risk of a mistake, and when a contract is voidable—create a possibility that rescission is improper. It is possible judges used their pre-existing knowledge of the full scope of contract doctrine to adjudicate the doctrinal question elements of rescission.

<sup>205</sup> In this loss condition, fourteen of the seventeen judges rescinded the contract; only three judges failed to rescind. *Id.*

<sup>206</sup> In the foregone gain condition, thirteen of the thirty-two judges rescinded; nineteen judges failed to rescind the contract. *Id.*

<sup>207</sup> As contract disputes heard by judges and international arbitrators likely vary, it was necessary to adjust the hypothetical to keep materials as realistic as possible within experimental constraints. Parties' natures can vary, contract subject matter varies, amounts in dispute are larger, and practices regarding expert valuation can vary. Given the transnational context, we did not rely on U.S. legal materials when instructing arbitrators on the applicable law.

wrong site); but in reality, LaKapa contained gold deposits and was worth US\$500 million. In a second version, the concession still involved “LaKapa,” but this time, both parties believed the site contained gold and agreed on a contract price of US\$380 million. A jointly appointed independent expert valued the site at US\$500 million but again mistakenly surveyed the wrong location; in reality, LaKapa contained significant iron pyrite deposits and was only worth US\$2.5 million. The arbitrators were also instructed that under the applicable law, parallel to the instruction in the videogame hypothetical, “a contract is voidable when ‘a mistake of both parties at the time a contract was made as to a basic assumption on which the contract was made has a material effect on the agreed exchange of promises.’”<sup>208</sup> Like the judges’ hypothetical, the contract dispute provided two frames, and in both conditions, parties made mutual mistakes about the contract.

Regardless of condition, 89.9% of international arbitrators rescinded the contract.<sup>209</sup> We nevertheless found some evidence of a small framing effect. Where a buyer sought rescission, a slightly smaller proportion of arbitrators (6.6%) disregarded the applicable law to enforce the contract. In contrast, where a seller sought rescission, a slightly larger proportion of arbitrators (14%) enforced the contract, as shown in Table 6.

Table 6: Contract Rescission and Framing: Percentage Rescinding  
(*n* = responses)

	Rescind Contract	Enforce Contract
<u>Intended Purchase of Gold Mine:</u> Buyer/Investor Seeks Rescission	93.4% (127)	6.6% (9)
<u>Intended Purchase of Fool’s Gold:</u> Seller/State Seeks Rescission	86% (104)	14% (17)
Total	89.9% (231)	10.1% (26)

Overall, the data demonstrate that framing appeared to influence international arbitrators.<sup>210</sup> As compared to their judicial counterparts answering

---

<sup>208</sup> As international arbitrators come from different legal traditions, our experiment omitted any reference to the *Restatement (Second) of Contracts* and provided a clean statement of the governing law.

<sup>209</sup> Two hundred thirty-one arbitrators rescinded the contract and twenty-six enforced. Five arbitrators did not respond.

<sup>210</sup> A Fisher’s exact test revealed that framing was marginally significant ( $p = 0.06$ ;  $n = 257$ ). The technical non-significance could be due to insufficient power. *Ex post* power analysis reveals that power of the analysis was 60%. *A priori* power analysis reveals a sample of 343—nearly 100 more arbitrators—would be required to reliably ascertain the lack of a framing effect. Although the Fisher’s test is arguably preferable, a Pearson’s Chi-Square Test of Independence revealed that arbitrators were reliably affected by whether the claimant was a buyer

a different question, arbitrators did not appear more susceptible to framing than the Utah judges. Framing could, however, operate both similarly and differently when comparing arbitrators and judges. One can imagine that it is possible that other types of judges, who answered the precise hypothetical we administered to arbitrators—rather than the hypothetical involving consumers, without an independent expert report, and with an express reference to the *Restatement (Second) of Contracts*<sup>211</sup>—could perform differently than the Utah judges and more similarly to the international arbitrators. The more critical insight, based upon our framing experiments, is that we were unable to identify evidence that the cognition of international arbitrators was inferior to their judicial counterparts.

As a final matter, although it is not a traditional cognitive illusion but is of interest to international arbitration, we tested whether rescission decisions varied based on whether the arbitrator was appointed by the claimant, respondent, or a neutral arbitration institution. The tests were unable, however, to ascertain a reliable link: (1) between appointment process and rescission decisions,<sup>212</sup> or (2) the larger model that tested interactions among framing and appointment variables.<sup>213</sup> Table 7 provides a breakdown of arbitrator responses purely as a function of appointment.

---

or seller ( $\chi^2(1) = 3.889$ ;  $p = 0.049$ ;  $r = 0.16$ ). Using Cohen's convention, the effect size for the significant effect was statistically small.

<sup>211</sup> See *supra* note 205.

<sup>212</sup> When focusing purely on the appointment variable, a Pearson's Chi-Square Test of Independence failed to confirm our hypothesis that appointment affected decisions ( $\chi^2(2) = 0.181$ ;  $p = 0.91$ ;  $r = 0.03$ ;  $n = 257$ ). The overall pattern was, irrespective of appointment condition roughly 90% of arbitrators correctly applied the applicable law and rescinded the contract. Inferences about the lack of a reliable relationship are improper as *ex post* power analysis reveals that, because of the small effect size, the power of the analysis was 30%–40%.

<sup>213</sup> For the 2x3 design that analyzed both the frame and the appointment conditions, it was not possible to identify an interaction where frame and appointment variations produced meaningfully different rescission decisions ( $\chi^2(5) = 8.121$ ;  $p = 0.15$ ;  $r = 0.18$ ;  $n = 257$ ). For buyer/investor claims: (a) with buyer/investor appointment, forty-two (89.4%) rescinded and five enforced; (b) with seller/state appointment, forty-two (95.5%) rescinded and two enforced; and (c) for ICSID appointment, forty-three (95.6%) rescinded and two enforced. For seller/state rescission: (a) with seller/state appointment, thirty-six (92.3%) rescinded and three enforced; (b) with buyer/investor appointment, twenty-nine (80.6%) rescinded and seven enforced; (c) with ICSID appointment, thirty-nine (84.8%) rescinded and seven enforced. For the 2x3 design, the power of the analysis was between 0.60–0.70. Although standard social science protocols tolerate an error of 20%, the *ex post* power analysis reflects a 30%–40% risk of error. Note 214, *infra*, offers an *a priori* power analysis of the sample required to reliably identify the reliable lack of an effect.

Table 7: Contract Rescission and Appointment: Percentage Rescinding  
( $n$  = responses)

	Rescind Contract	Enforce Contract
Claimant Appointment	90.7% (78)	9.3% (8)
Respondent Appointment	88.8% (71)	11.2% (9)
Institutional Appointment	90.1% (82)	9.9% (9)
Total	89.9% (231)	10.1% (26)

We observe that, as these are null results (i.e., the tests were unable to identify a statistically significant effect), particular caution is warranted in drawing inferences. The null results cannot prove the lack of an effect, but they are a piece of evidence to consider before drawing firm conclusions about the intuitive predisposition of international arbitrators.<sup>214</sup> We nevertheless offer the raw data in the hopes of advancing the science of arbitrator decisionmaking.<sup>215</sup>

#### D. Representativeness

People often make evaluations based on surface similarities, rather than base rates or statistical realities. One archetypal example of this phenomenon, which psychologists call “representativeness,”<sup>216</sup> can be illustrated by the classic case,

---

<sup>214</sup> There are a variety of reasons to be cautious about drawing strong inferences from the results. For instance, for a 2x3 design to have acceptable power, a sample of 1029 arbitrators would be required. This necessitates testing replication would require over 700 additional arbitrators to make a reliable conclusion about the lack of a statistical effect. Relatedly, there are concerns about eternal validity. For example, the two to three years it may take to resolve a case, the potential financial self-interest in repeat appointment, and the implications of interactions with co-arbitrators could mean our study was unrealistic on appointment-related matters; and inferences drawn from a hypothetical on appointment in this experimental setting are limited.

<sup>215</sup> We observe, for example, that appointment effects might be constrained by clear rules of law and minimal arbitrator discretion. One preliminary study suggested that, in one limited situation, appointment could influence outcomes; namely, where an arbitrator made a decision on costs, a winning party-appointed arbitrator (possibly appointed by either an investor or a state) often made a 100% cost shift in favor of the winner. Sergio Puig & Anton Strezhnev, *Affiliation Bias in Arbitration: An Experimental Approach* 24–25 (Ariz. Legal Studies, Discussion Paper No. 16-31) (copy on file with author). Puig and Strezhnev’s research may, however, be confounded by the failure to address that successful investors reliably have costs shifted in their favor but successful states did not. The research nevertheless raises the possibility that, in areas of arbitral discretion, an “appointment effect” might contribute to arbitral decisionmaking; but likewise, where there is clear law and bounded discretion, there could be decreased risk. Future research should explore this in greater detail.

<sup>216</sup> Daniel Kahneman & Amos Tversky, *Subjective Probability: A Judgment of Representativeness*, 3 COGNITIVE PSYCHOL. 430 (1972); Amos Tversky & Daniel Kahneman, *Judgments of Representativeness*, in JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 84, 84–85 (Daniel Kahneman, Paul Slovic & Amos Tversky eds., 1982).

*Byrne v. Boadle*.<sup>217</sup> Previous research has used that case to test cognitive illusions in judicial decisionmaking:

The plaintiff was passing by a warehouse owned by the defendant when he was struck by a barrel, resulting in severe injuries. At the time, the barrel was in the final stages of being hoisted from the ground and loaded into the warehouse. The defendant's employees are not sure how the barrel broke loose and fell, but they agree that either the barrel was negligently secured or the rope was faulty. Government safety inspectors conducted an investigation of the warehouse and determined that in this warehouse: (1) when barrels are negligently secured, there is a 90% chance that they will break loose; (2) when barrels are safely secured, they break loose only 1% of the time; (3) workers negligently secure barrels only 1 in 1,000 times.<sup>218</sup>

In the earlier scholarship, researchers then asked the generalist judges: "Given these facts, how likely is it that the barrel that hit the plaintiff fell due to the negligence of one of the workers?" Judges could then select one of four options: (a) 0%–25%, (b) 26%–50%, (c) 51%–75%, or (d) 76%–100%.<sup>219</sup> The correct answer, which requires deductive analysis, reveals the correct probability of negligence was 8.3%.<sup>220</sup> A more intuitive response, however, would treat the 90% figure as the likelihood that negligence caused the accident, thereby converting the 90% statistic (i.e., the likelihood of negligence) into something else (i.e., the likelihood of negligence given the injury).<sup>221</sup>

The earlier research observed that judges used intuitive, representative thinking rather than rational, deductive thought. They found that only 40.9% of judges selected the correct response, while a comparable 40.3% selected the

---

<sup>217</sup> (1863) 159 Eng. Rep. 299 (Ex. Ch.).

<sup>218</sup> Guthrie, Rachlinski & Wistrich, *supra* note 10, at 808 (quoting *Byrne*, 159 Eng. Rep. 299).

<sup>219</sup> *Id.*; Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 22–23.

<sup>220</sup> See Guthrie, Rachlinski & Wistrich, *supra* note 10, at 809 ("Because the defendant is negligent .1% of the time and is 90% likely to cause an injury under these circumstances, the probability that a victim would be injured by the defendant's negligence is .09% (and the probability that the defendant is negligent but causes no injury is .01%). Because the defendant is not negligent 99.9% of the time and is 1% likely to cause an injury under these circumstances, the probability that on any given occasion a victim would be injured even though the defendant took reasonable care is 0.999% (and the probability that the defendant is not negligent and causes no injury is 98.901%). As a result, the conditional probability that the defendant is negligent given that the plaintiff is injured equals .090% divided by 1.089%, or 8.3%."); see also Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 23 n.125.

<sup>221</sup> Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 22–23; Guthrie, Rachlinski & Wistrich, *supra* note 10, at 808–10; see also Jeffrey J. Rachlinski, *Bottom-Up Versus Top-Down Lawmaking*, 73 U. CHI. L. REV. 933, 939 (2006).

intuitive, but incorrect, response, professing the belief that the accident was more than 75% likely to have resulted from negligence.<sup>222</sup>

We gave the arbitrators who participated in our study a nearly identical question, involving a storage company that had contracted with an engineering corporation to deliver and maintain storage equipment. After a product defect caused damage, the storage company pursued arbitration to recover for breach of contract warranties and negligent equipment maintenance. The arbitrators learned that government safety inspectors:

determined that in this warehouse: (a) when containers are negligently secured, there is a 90% chance that they will break loose; (b) when containers are safely secured, they break loose only 1% of the time; (c) workers negligently secure containers only 1 in 1,000 times. Given these facts, how likely is it that the container fell due to the negligence of one of Storage's workers?

We asked participants to select one of four probability ranges: "(a) 0–25%, (b) 26–50%, (c) 51–75%, or (d) 76–100%." Of responding arbitrators,<sup>223</sup> 60.6% ( $n = 152$ ) answered the question correctly, while only 17.9% ( $n = 45$ ) of the international arbitrators selected the erroneous, but intuitive, answer. A small proportion of international arbitrators selected incorrect non-intuitive answers, with 8.8% ( $n = 23$ ) and 11.8% ( $n = 31$ ) selecting options (b) and (c), respectively.<sup>224</sup> On a structurally similar problem, international arbitrators outperformed domestic judges, as Figure 3 suggests.<sup>225</sup>

---

<sup>222</sup> Guthrie, Rachlinski & Wistrich, *Blinking*, *supra* note 77, at 23–24; Guthrie, Rachlinski & Wistrich, *supra* note 10, at 809–10.

<sup>223</sup> All participants received this hypothetical. Eleven arbitrators (4.2%) failed to respond.

<sup>224</sup> Thirty-two arbitrators (12.4%) made manuscript comments to calculate probabilities.

<sup>225</sup> A Fisher's exact test compared the correct and incorrect assessments of international arbitrators and U.S. federal magistrate judges. The test demonstrated arbitrators were reliably better at identifying the correct answer ( $p = 0.0001$ ;  $r = 0.19$ ;  $n = 410$ ). Whereas 152 arbitrators (60.6%) answered correctly and 99 answered incorrectly, 65 judges (40.9%) answered correctly and 94 answered incorrectly. Guthrie, Rachlinski & Wistrich, *supra* note 10, at 809–10; Figure 3.



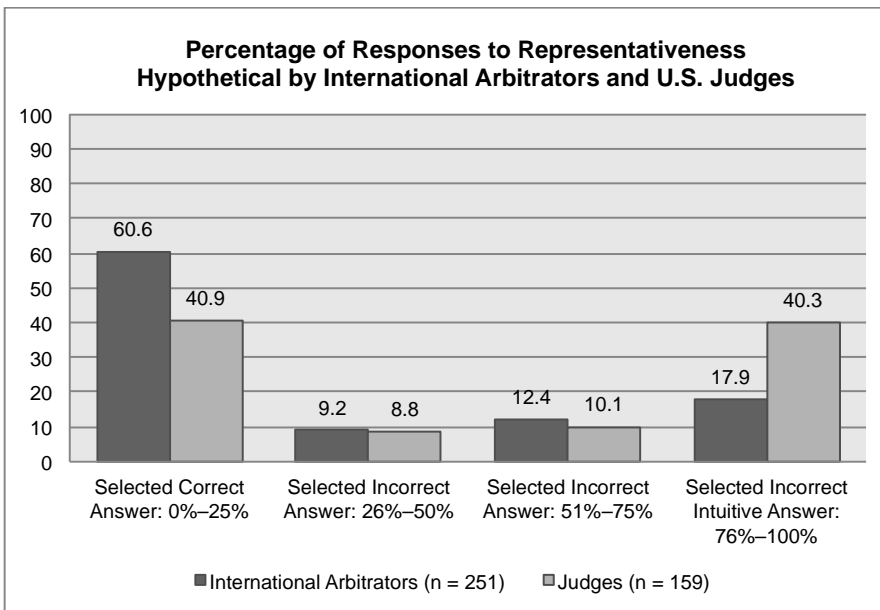


Figure 3. Percentage of Responses to Representativeness Hypothetical by International Arbitrators and U.S. Judges.

Arbitrators also facially outperformed other professionals responding to similar questions. In a *New England Journal of Medicine* study, fewer than 20% of doctors at Harvard teaching hospitals responding to a similar problem answered correctly, while 45% provided intuitive and incorrect responses.<sup>226</sup> Only 10% of Norwegian law students correctly answered a similar problem, while 58% of students selected the intuitive response.<sup>227</sup> In short, international arbitrators, though influenced by representativeness, appeared to do somewhat better than other professionals confronting similar problems.

<sup>226</sup> See Ward Casscells, Arno Schoenberger & Thomas B. Graboys, *Interpretation by Physicians of Clinical Laboratory Results*, 299 *NEW ENG. J. MED.* 999, 1000 (1978) (stating “[e]leven of the 60 participants, or 18 per cent, gave the correct answer” and noting twenty-seven subjects (45%) selected the intuitive, incorrect response).

<sup>227</sup> Erling Eide, *Two Tests of Base Rate Neglect Among Law Students* (2011), [http://www.uio.no/studier/emner/jus/jus/JUS4121/v12/undervisningsmateriale/Evidence\\_RLE2\\_kopi\\_4\\_avd.pdf](http://www.uio.no/studier/emner/jus/jus/JUS4121/v12/undervisningsmateriale/Evidence_RLE2_kopi_4_avd.pdf). The sampled Norwegian law students may differ from law students elsewhere. When beta-testing the entire first-year class at Washington & Lee Law School in January 2014 using the same question administered to arbitrators, 54% ( $n = 54$ ) selected the correct answer and 11% ( $n = 11$ ) selected the intuitive incorrect answer.

### E. Egocentrism

People routinely overestimate their talents and life prospects. In one classic study, for example, researchers found that recently married U.S. couples almost unanimously expected they would not divorce, even though they knew the divorce rate was 50%.<sup>228</sup> Psychologists call this phenomenon the egocentric or self-serving bias.

Egocentric bias can negatively impact adjudicators and the parties who appear in front of them because it can “prevent judges from maintaining an awareness of their limitations . . . [and] may make it hard for judges to recognize that they can and do make mistakes.”<sup>229</sup>

Researchers have explored whether egocentrism affects judges. Eisenberg originally identified a self-serving bias in bankruptcy judges. Specifically, judges evaluated themselves as more fair, more efficient, and better case managers than lawyers evaluated the same judges.<sup>230</sup> Guthrie, Rachlinski, and Wistrich similarly explored whether generalist<sup>231</sup> and specialist<sup>232</sup> judges exhibited egocentrism. They found U.S. adjudicators likewise had self-serving views of their adjudicative skills.

To explore whether and how egocentrism affects arbitrators, we randomly gave subjects two questions asking them to assess themselves on several arbitrator tasks—i.e., assessing witness credibility, making quality decisions, providing parties with procedural efficiency, and the rate of challenges to their awards. We asked the arbitrators to place themselves in one of four quartiles: (1)

---

<sup>228</sup> Lynn A. Baker & Robert E. Emery, *When Every Relationship Is Above Average: Perceptions and Expectations of Divorce at Time of Marriage*, 17 LAW & HUM. BEHAV. 439, 443 (1993); cf. Ola Svenson, *Are We All Less Risky and More Skillful than Our Fellow Drivers*, 47 ACTA PSYCHOLOGICA 143, 146 (1981) (finding 88% of U.S. drivers and 77% of Swedish drivers believed themselves to be safer driver's than the median, but observing more U.S. drivers (46.3%) placed themselves in the most skilled group as compared to Swedes (15.5%)).

<sup>229</sup> Guthrie, Rachlinski & Wistrich, *supra* note 10, at 815.

<sup>230</sup> Theodore Eisenberg, *Differing Perceptions of Attorney Fees in Bankruptcy Cases*, 72 WASH. U. L.Q. 979, 982 (1994). While 96% of judges reported ruling on requests for interim awards within thirty days, only 79% of lawyers reported that judicial conduct. *Id.* at 984. Compared to lawyers' assessments, bankruptcy judges perceived themselves as more closely monitoring cases and providing efficient fee reimbursement. *Id.* at 984–87. See generally Jane Goodman-Delahunty et al., *Insightful or Wishful: Lawyers' Ability to Predict Case Outcomes*, 16 PSYCHOL. PUB. POL'Y & L. 133 (2010).

<sup>231</sup> Guthrie, Rachlinski & Wistrich, *supra* note 10.

<sup>232</sup> Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1519–20.

the top 25%, (2) the second quartile, (3) the third quartile, or (4) the bottom 25%.<sup>233</sup>

As shown in Table 8, we found that international arbitrators, like U.S. adjudicators, provided egocentric or self-serving interpretations of their adjudicative skills. The distribution of arbitrators' responses departs significantly from what one would expect if there was no egocentrism bias.<sup>234</sup>

Table 8: Self-assessment of Adjudicative Skill: International Arbitrators, Judges, and ALJs.<sup>235</sup>

Skill (sample)	Percentile in each Quartile			
	Top Quartile: Better than 75%	Second Quartile: Better than 50%, worse than top 25%	Third Quartile: Better than 25%, worse than top 50%	Bottom Quartile: Worse than 75%
Evaluating Credibility International				
Arbitrators (124):	29.8%	46.8%	21.0%	2.4%
ALJs (36):	25.0%	58.0%	17.0%	0.0%
Unbiased or Impartial Decisions: International				
Arbitrators (124):	43.5%	41.1%	13.7%	1.6%
ALJs (36):	50.0%	47.0%	3.0%	0.0%
Procedural Efficiency: International				
Arbitrators (124):	36.3%	55.6%	7.3%	0.8%
Challenges to Decisions: International				
Arbitrators (122):	53.3%	32.8%	10.7%	3.3%
Judges (155):	56.1%	31.6%	7.7%	4.5%

In response to the query, "If the researchers were to rank all of the arbitrators currently in this room according to their skill at reliably assessing witness

<sup>233</sup> The instruction was to evaluate, based upon those in the room, whether arbitrators fell in the highest, second highest, second lowest, or lowest quartile for a specific skill. See *infra* note 239.

<sup>234</sup> In the absence of egocentrism, results should have been evenly distributed across quartiles. Instead, there was a large and meaningful departure for responses to questions on witness credibility ( $t(123) = 27.983$ ;  $p < 0.001$ ;  $r = 0.93$ ;  $n = 124$ ), efficiency in dispute resolution ( $t(123) = 30.549$ ;  $p < 0.001$ ,  $r = 0.94$ ;  $n = 124$ ), impartial decisionmaking ( $t(123) = 25.554$ ;  $p < 0.001$ ,  $r = 0.92$ ;  $n = 124$ ), and challenges to awards ( $t(121) = 22.534$ ;  $p < 0.001$ ,  $r = 0.90$ ;  $n = 122$ ). See COHEN, *supra* note 135, 113–16 (noting Cohen's conventions that a "large" effect is present when  $r \geq 0.50$ ).

<sup>235</sup> Data from Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1519–20.

credibility, what would your rate be?”, 76.6% of arbitrators<sup>236</sup> identified that they were better than the median arbitrator in the room.

Arbitrators were even more bullish in assessing their capacity to provide unbiased decisions. When asked: “If the researchers were to rank all of the arbitrators currently in this room according to their skill at making accurate and impartial decisions, what would your rate be?”, nearly 85% of responding arbitrators<sup>237</sup> indicated they were better than the median arbitrator present at an elite conference.

Arbitrators were most self-serving when assessing their procedural efficiency. When asked, “If the researchers were to rank all of the arbitrators currently in this room according to their skill at efficiently resolving disputes in a timely manner, what would your rate be?”, nearly 92% of arbitrators<sup>238</sup> ranked themselves as superior to the median arbitrator in attendance.

In addition to evaluating themselves more favorably than their counterparts, international arbitrators also assumed their decisions had been challenged much less frequently than the decisions of their peers. We asked: “If the researchers were to rank all of the arbitrators currently in this room according to the rate at which their decisions have been challenged during their careers, what would your rate be?”<sup>239</sup> For the international arbitrators, 86.1% assumed that their reversal rates were better (i.e., lower rates of challenge) than the median arbitrator in the room.

Direct comparisons between judges and arbitrators remain difficult. On balance, we found arbitrators and judges were similarly influenced by

---

<sup>236</sup> One hundred thirty-one arbitrators received the credibility question. Seven arbitrators (5.3%) failed to answer.

<sup>237</sup> One hundred twenty-nine arbitrators received the decision-making question. Five arbitrators (3.9%) did not answer.

<sup>238</sup> One hundred thirty-three arbitrators received the efficiency question, but nine (6.7%) failed to answer.

<sup>239</sup> This information can be found in the stimulus materials, which are available upon request from the lead author.

egocentrism,<sup>240</sup> though international arbitrators seemed somewhat less likely than ALJs to overestimate the quality and integrity of their decisions.<sup>241</sup>

#### IV. INTERPRETATION AND IMPLICATIONS

In this first-ever psychological experiment involving international arbitration, we found that arbitrators often made intuitive and impressionistic decisions rather than the fully rational and deliberative decisions that might be normatively desirable. This finding, though perhaps disappointing, is unsurprising. Arbitrators are people, and they make judgments and decisions the way other people do.

More encouraging, and perhaps more surprising, we found evidence contradicting common tropes against arbitration. Although direct comparisons are difficult, our data revealed that international arbitrators performed at least as well as, but never demonstrably worse than, judges. There may be reasons to prefer judges to arbitrators,<sup>242</sup> but quality of judgment and decisionmaking, at

---

<sup>240</sup> A Fisher's Exact Test was unable to identify a meaningful difference in the responses of international arbitrators and ALJs self-assessments of capacity to evaluate witness credibility. Using ALJ responses from Guthrie, Rachlinski & Wistrich, *Hidden Judiciary*, *supra* note 77, at 1520, it was not possible to detect different response patterns for ALJs ranking themselves above ( $n = 30$ ) or below the median ( $n = 6$ ) and arbitrators who ranked themselves above ( $n = 95$ ) and below ( $n = 29$ ) the median ( $p = 0.495$ ;  $r = 0.07$ ;  $n = 160$ ). Similarly, a Fisher's exact test was unable to detect a meaningful difference in how magistrate judges and international arbitrators self-assessed whether their decisions would be successfully challenged in later court action ( $p = 0.72$ ;  $r = 0.02$ ;  $n = 277$ ). Guthrie, Rachlinski & Wistrich, *supra* note 10, at 809–10, found 136 judges ranked themselves above the median in reversal rates (i.e., having low reversal rates), whereas nineteen judges ranked themselves as being below the median. One hundred five arbitrators ranked themselves as being superior to the median (i.e., having lower challenge rates) and seventeen arbitrators evaluated themselves as being in the two lowest quartiles (i.e., having a higher challenge rate). The results lacked sufficient power to definitively exclude presence of a relationship; but as the effect size was less than small, null-results may not reflect low power.

<sup>241</sup> A Fisher's exact test analyzed ALJs ranking themselves above ( $n = 35$ ) and below ( $n = 1$ ) the median for unbiased decisions, and international arbitrators ranking themselves above ( $n = 105$ ) and below ( $n = 19$ ) the median for impartial decisions. A greater proportion of ALJs over-estimated their skill in making unbiased decisions as compared to arbitrators ( $p = 0.048$ ,  $r = 0.16$ ). This does not mean arbitrators were immune from egocentrism in evaluating their capacity to make impartial decisions, as data reflects they fell prey to the same fallacy. Rather, a lower proportion of international arbitrators self-identified high skills and greater proportion were somewhat more modest. Slight variations in wording limit the value of comparison, however.

<sup>242</sup> There may be legitimate concerns about arbitration, including concerns of public access and transparency. See Owen M. Fiss, *Against Settlement*, 93 YALE L.J. 1073, 1075–76, 1078 (1984). UNCITRAL's new treaty and arbitration rules provide increased transparency, particularly in ITA. G.A. Res. 69/116, United Nations Convention on Transparency in Treaty-Based Investor-State Arbitration (Dec. 10, 2014), <http://www.uncitral.org/pdf/english/texts/arbitration/transparency-convention/Transparency-Convention-e.pdf>; G.A. Res. 68/109, UNCITRAL Rules on Transparency in Treaty-Based Investor-State Arbitration (Dec. 16, 2013), <http://www.uncitral.org/pdf/english/texts/arbitration/rules-on-transparency/Rules-on-Transparency-E.pdf>. Concerns about arbitrators' incentives for ethical conduct can and should be addressed. Existing duties of impartiality and laws permit challenge and dismissal of biased arbitrators. See *supra* notes 70 and accompanying

least as measured in these experimental studies, is not one of them. In addition, our work casts doubt on the common assumption that arbitrators simply “split the baby” when making decisions. Rather, confirming existing research on real awards,<sup>243</sup> international arbitrators did not appear to “split the baby” when making awards.<sup>244</sup>

The experimental results explicate the possible ongoing normative value of arbitration. Having failed to identify international arbitrators' inferior capacity, the possibility remains that structural safeguards might improve arbitration and to decrease risk of error. We acknowledge the limitations of the results yet wish to make normative recommendations that system designers may consider for managing disputes.

#### A. *Limitations*

The fact that we find evidence of intuitive decisionmaking in our experimental research does not conclusively demonstrate that arbitrators behave similarly during actual proceedings.

First, as we have noted in this Article and elsewhere,<sup>245</sup> selection effects limit the value of inferences. Because we do not know, and likely will never be able to know, the demographic characteristics of the global population of international arbitrators, we cannot definitively confirm how representative our sample might be. It is possible that the international arbitrators who attended ICCA and participated in our study skewed older, more economically advantaged, more elite, and with a greater proportion of women.<sup>246</sup>

---

text. The recently signed Trans-Pacific Partnership includes a “code of conduct” for international arbitrators. OFFICE OF THE U.S. TRADE REPRESENTATIVE, SUMMARY OF THE TRANS-PACIFIC PARTNERSHIP AGREEMENT (2015), <https://ustr.gov/about-us/policy-offices/press-office/press-releases/2015/october/summary-trans-pacific-partnership>. A full discussion of net normative costs and benefits of international arbitration is beyond this Article, which focuses on experimental manipulation in search of evidence-based insights for targeted reform and informed decisionmaking. *See supra* notes 4, 58 (describing concerns about international arbitration addressed by other literature).

<sup>243</sup> *Supra* note 152 and accompanying text.

<sup>244</sup> In contradiction to claims that arbitrators are intuitively predisposed to parties appointing them, our experiment was unable to identify evidence that party appointment reliably influenced contract rescission. *See supra* notes 213–14 and accompanying text. These null results also come with limitations. *See supra* note 215.

<sup>245</sup> *See, e.g., supra* notes 133–34, 165, 244–46 (identifying some of the issue-specific limitations); *see also* Franck et al., *supra* note 4, at 443–46, 501.

<sup>246</sup> Franck et al., *supra* note 4, at 443–45. Recent research suggests female arbitrators remain less than 10% of the population of international arbitrators. Lucy Greenwood & C. Mark Baker, *Is the Balance Getting Better? An Update on the Issue of Gender Diversity in International Arbitration*, 31 *ARB. INT'L* 413, 415 (2015). By

Second, international arbitrators' conduct in real disputes could differ from responses to our hypotheticals. International arbitration proceedings are often lengthy, complex, and rely upon numerous witness statements and voluminous documents. Rather than making snap judgments during a survey, arbitrators have access to time, resources, tribunal secretaries who function like judicial clerks, and group deliberations.<sup>247</sup> Applicable substantive and procedural law could inject debiasing mechanisms that limit the influence of cognitive illusions. The differences between our experiment and the natural ecology of international arbitration therefore necessitate caution in making strong inferences. Nevertheless, the results likely have some application beyond the laboratory. The research employed standard cognitive psychology research methods used successfully on other adjudicators for over a decade, and similar methodologies have proven successful in identifying strategies people use to make decisions in real life.<sup>248</sup> Moreover, decisionmakers may be even more likely to rely on cognitive shortcuts, like anchoring and framing, in real-world settings precisely because of the volume of information and complexity of decisionmaking.

Third, the inherent heterogeneity of international economic disputes may limit the external validity of our research. Our materials explored decisionmaking within ICA and ITA disputes. It is possible the results are not generalizable more broadly given variation in disputes, context, facts, applicable law, or culture.

Fourth, several hypotheticals asked arbitrators to make independent decisions, as if they were a sole arbitrator or one arbitrator on a three-member tribunal.<sup>249</sup> It is unusual for arbitrators to act alone; rather, three-member tribunals are standard. Our research was unable to capture one of the fundamental rule-of-law values and debiasing tools embedded within

---

contrast, our study identified roughly 17% of the arbitrators were female, which creates a risk women were overrepresented in our research.

<sup>247</sup> Group deliberation could, but need not, guarantee enhanced quality. *Infra* note 262.

<sup>248</sup> Guthrie, Rachlinski & Wistrich, *supra* note 10, at 819; Daniel Kahneman & Amos Tversky, *On the Reality of Cognitive Illusions*, 103 *PSYCHOL. REV.* 582, 582 (1996). Others dispute the influence of cognitive psychology. Gerd Gigerenzer, *How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases"*, 2 *EUR. REV. SOC. PSYCHOL.* 83, 84–85, 109–10 (1991).

<sup>249</sup> In the one hypothetical that manipulated party appointment—and placed subjects in the role of acting as a claimant, respondent, or institutional appointee—we were unable to identify that appointment reliably affected arbitrators' legal decisions. *See supra* note 213. There is a difference, however, between answering a hypothetical question during a thirty to forty minute survey and living through a case for two to three years as a party-appointee. While we may have captured some aspects of arbitrator intuition, this does not address the sustained influence of environmental factors occurring over an arbitration's lifetime.

international arbitration—namely, the deliberation process. Future research should explore panel effects in decisionmaking.

Fifth, the results may be *sui generis* to international arbitration. International dispute settlement is highly specialized, involving different practices, different legal rules, and high barriers to entry.<sup>250</sup> It is possible that inferences for domestic arbitration are limited, as international arbitrators do not necessarily adjudicate domestic consumer and employment disputes. Likewise, although there is a degree of overlap between international arbitrators and judges on international courts and tribunals,<sup>251</sup> that overlap is not complete. Caution is therefore warranted about the scope of inferences.

Sixth, we acknowledge that linguistic capacity could influence responses. Yet, English is a *lingua franca*<sup>252</sup> and has become dominant in international arbitration.<sup>253</sup> It is possible that a conference in English did not generate a large selection effect, as those without English skills may not be actively engaged in international arbitration. Yet the risk is not eliminated, and non-native English speakers could systemically differ. Native language, however, did not explain the performance of international arbitrators. For those items where arbitrators performed particularly well, we were unable to identify meaningful differences in the responses of native and non-native English speakers.<sup>254</sup>

---

<sup>250</sup> DEZALAY & GARTH, *supra* note 81, at 10, 124, 198; Rogers, *Vocation*, *supra* note 3, at 963–64.

<sup>251</sup> Past and present judges on the International Court of Justice (ICJ) have been arbitrators. Multiple ICJ members have been ITA arbitrators in disputes or *ad hoc* committees, including: James Crawford, Chris Greenwood, Peter Tomka, Joan Donoghue, Abdulqawi Ahmed Yusuf, and Patrick Robinson. Two former ICJ judges (Bruno Simma and Stephen Schwebel) were also arbitrators. Charles Brower and David Caron were or are serving as a judge on the Iran-U.S. Claims Tribunal; and both have been international arbitrators. Giorgio Sacerdoti, Georges Abi-Saab, Florentino Feliciano, and Donald McRae have been arbitrators and WTO adjudicators. See José Augusto Fontoura Costa, *Comparing WTO Panelists and ICSID Arbitrators*, 1 ONATI SOCIO-LEGAL SERIES, 2011, at 1, 14; Joost Pauwelyn, *Rule of Law Without the Rule of Lawyers? Why Investment Arbitrators Are from Mars, Trade Adjudicators from Venus*, 109 AM. J. INT'L L. 761, 768–69 (2015).

<sup>252</sup> Timothy Lau, *Offensive Use of Prior Art to Invalidate Patents in U.S. and Chinese Patent Litigation*, 30 UCLA PAC. BASIN L.J. 201, 250 (2013).

<sup>253</sup> Roger P. Alford, *The American Influence on International Arbitration*, 19 OHIO ST. J. ON DISP. RESOL. 69, 86 (2003); Stephan W. Schill, *W(h)ither Fragmentation? On the Literature and Sociology of International Investment Law*, 22 EUR. J. INT'L L. 875, 887 (2001).

<sup>254</sup> Bivariate correlations could not identify reliable links between native English speakers or non-native speakers for: CRT scores ( $r(232) = 0.08$ ;  $p = 0.24$ ) or correct responses on representativeness ( $r(243) = -0.03$ ;  $p = 0.61$ ). Native English capacity was not reliably associated with responses on the beach front property ( $r(95) = 0.08$ ;  $p = 0.43$ ) or contract rescission ( $r(249) = -0.05$ ;  $p = 0.41$ ) hypotheticals. As all analyses were less than statistically small ( $r < 0.10$ ), the analysis may not be underpowered. A sample of 781 arbitrators would be sufficiently powered to reliably exclude the possibility of a native-language effect.



Lastly, in those instances when tests were unable to detect differences, it is not possible to rule differences out. Although many latent effects were statistically small, various tests were statistically underpowered. *Post hoc* power analyses revealed samples of 780–1000 arbitrators would be required to ascertain the lack of an effect.<sup>255</sup> More research is therefore necessary. We acknowledge the limitations and hope this first-generation research provides a baseline for future scholarship.

## *B. Normative Implications*

In this study, we sought to explore whether international arbitrators, like other adjudicators, make decisions in ways that depart from the rational actor model in transnational adjudication. We found that they do, and this, in turn, has implications for the design of transnational dispute settlement systems (as well as domestic dispute systems).

### *1. Allocating Adjudicative Authority*

Our research demonstrates that, like other expert adjudicators, international arbitrators were susceptible to cognitive illusions including anchoring, framing effects, representativeness, and egocentrism. While comparisons between arbitrators and judges are difficult, we found no evidence that arbitrators were inferior to judges.

When choosing to allocate adjudicative authority to judges or arbitrators, then, designers of dispute settlement systems should not presume that national (or international) judges will inevitably provide decisionmaking services that are superior to international arbitrators. Earlier research demonstrates that national judges over-rely on intuition and make errors in legal decisionmaking. So do arbitrators. Regardless of title or mandate, adjudicators are fallible beings who generate error, even with the best of intentions and effort. When making normative design choices about who should resolve international disputes, arbitrators should neither be favored *nor* disfavored based on their cognitive skill.

While there are undoubtedly costs to using arbitration—including paying for the services of the arbitrators and related institutions—there are likewise benefits. International arbitration might offer opportunities to minimize the risk

---

<sup>255</sup> See, e.g., *supra* notes 133, 164, 210, 212–14 (offering power analyses and identifying requisite sample size).

of some adjudicative error. For example, international arbitrators are a transnational group speaking multiple languages. Research suggests that people are less likely to fall prey to cognitive illusions when evaluating options in a language other than their mother tongue.<sup>256</sup> This generates the possibility that international arbitrators' standard practice, involving regular interaction with different languages or cultures, increases the likelihood of careful attention and focused deliberation. Beyond this, the elite nature of international arbitration—and strong barriers to entry—likely generate market forces where highly qualified lawyers and transnational professionals ultimately serve as arbitrators.<sup>257</sup>

## 2. *Minimizing Risk of Error Through Structure and Procedure*

If intuition influences all adjudicators, critics are correct that judicial systems require procedural safeguards to minimize inaccurate or sub-optimal decisionmaking. There is obviously a potential trade-off between accuracy on the one hand and speed on the other. Adding procedures or implementing other debiasing mechanisms could decrease the risk of error but increase the time and cost of a dispute settlement process. Thus, dispute system designers, as well as parties selecting among dispute resolution mechanisms, should carefully weigh the benefits and costs of procedures designed to enhance decision accuracy.<sup>258</sup>

International arbitration typically includes some structures that might have debiasing effects and could easily incorporate others.<sup>259</sup>

For example, parties in international arbitration typically choose to have disputes resolved by three-member tribunals. The presence of multiple

---

<sup>256</sup> See Micheline Favreau & Norman S. Segalowitz, *Automatic and Controlled Processes in the First- and Second-Language Reading of Fluent Bilinguals*, 11 *MEMORY & COGNITION* 565, 567 (1983) (theorizing foreign language evaluations require more deliberate processing and fewer intuitive assessments); Boaz Keysar, Sayuri L. Hayakawa & Sun Gyu An, *The Foreign-Language Effect: Thinking in a Foreign Tongue Reduces Decision Biases*, 23 *PSYCHOL. SCI.* 661, 661, 667 (2012) (observing framing effects and loss aversion disappeared or decreased when subjects were tested in a foreign language).

<sup>257</sup> It is difficult to make uniform observations about the over 190 national judiciaries. Some judges are elected or partisan. Others may be elite professionals, but lack linguistic and inter-cultural competencies. There may also be partisanship concerns deriving from national or regional sympathies. International judges may share characteristics of international arbitrators, including language skills, training in multiple legal systems, and inter-cultural competencies.

<sup>258</sup> Elm, *supra* note 83, 114–24 (proposing several amendments to the UNCITRAL Rules to debias arbitrators).

<sup>259</sup> Both international arbitration and litigation permit evidence testing. In arbitration, parties challenge material facts and applicable law, providing an opportunity to disrupt and assess claims rather than relying on intuition or supposition. Court litigation can be similar.

adjudicators, who must interact and deliberate together, permits coordinated decisionmaking and group deliberation. While group deliberation is not necessarily an unmitigated good—indeed, it can lead to polarization and exacerbate decision errors—it can facilitate collaborative deliberation that serves as a check upon intuitive assessments.<sup>260</sup>

Moreover, many, if not most, international arbitration panels produce written opinions. In ITA, for example, 100-plus page opinions are common. The process of opinion writing itself could also serve as a check on intuition and facilitate deliberation, leading to higher quality outcomes. In those instances where arbitrators are not required by governing rules to write opinions, parties could contract for opinions, if so inclined. Parties also might mandate that tribunals include subsections in awards, follow prescribed checklists, or provide substantive reasoning of critical, replicable issues, if the parties believe they would benefit from a more detailed and precise explication of decisionmaking.

This speaks to a general virtue of international arbitration. In arbitration, in contrast to litigation, parties can adopt procedural rules and structures to enhance adjudicative quality and minimize the risk of decision error. For example, parties can structure procedures to give arbitrators more time to devote to deliberation. Likewise, parties can draft arbitration agreements to inject additional procedural rigor to decrease risks of error from intuitive adjudication.

Consider, for example, anchors. Our results show that irrelevant anchors, particularly large anchors, influenced the damage assessments of international arbitrators. Parties to arbitrations might craft rules to require a good-faith pleading rule or to provide clear cost-shifting rules to incentivize accurate damage assessments at the start of the case.<sup>261</sup> Alternatively, rules might require parties to produce supported damage assessments at the outset so that, when pleading damages, parties do so with particularity and support. Other

---

<sup>260</sup> Group deliberations do not necessarily enhance quality or accuracy. DENNIS J. DEVINE, *JURY DECISION MAKING: THE STATE OF THE SCIENCE* 152–53, 158–59 (2012); Daniel Gigone & Reid Hastie, *Proper Analysis of the Accuracy of Group Judgments*, 121 *PSYCHOL. BULL.* 149, 149 (1997); Dan Simon, *More Problems with Criminal Trials: The Limited Effectiveness of Legal Mechanisms*, 75 *L. & CONTEMP. PROBS.*, 2012, No. 2, at 167, 193–200; Adrian Vermeule, *Many-Minds Arguments in Legal Theory*, 1 *J. LEGAL ANALYSIS* 1, 26–35 (2009).

<sup>261</sup> The ICSID Convention does not have clear cost-shifting rules, and tribunals have not offered consistent rulings on costs or a clear set of incentives for cost assessments. *See, e.g.*, Franck, *supra* note 11, at 801 & n.170 (noting that “there is no international convention on the treatment of costs in investment treaty arbitration”); David Smith, *Shifting Sands: Cost-and-Fee Allocation in International Investment Arbitration*, 51 *VA. J. INT’L L.* 749, 751–52 (2011) (noting that tribunals have rendered “scattershot” rulings on costs under the ICSID Convention).

stakeholders may wish to take this approach. For example, international arbitration institutions revising rules, national legislatures amending law, or states negotiating treaties may wish to explore providing clear default rules for international arbitration that: (a) require parties to plead damages with specificity and in good faith or (b) create incentives for reasonable, evidence-based damage claims. Such an approach could maximize the value of relevant anchors and decrease the risk that irrelevant anchors exert improper influence on damage assessments.

With each of these procedural innovations, there is tension between efficiency and accuracy. The tension may be more theoretical than real. One could imagine a complex international dispute, being decided by three subject matter experts, with adequate time and prepared parties, creating a written opinion that would minimize the chance for error and produce a just and timely result.<sup>262</sup> As long as parties and arbitrators acknowledge that humans must test intuition with deliberation, arbitration's flexibility allows parties and arbitrators to create a tailor-made process to do so.

#### CONCLUSION

Arbitrators, like judges, are fallible. Arbitrators, like judges, make intuitive decisions that they might, or might not, override with deliberation. Arbitrators, like judges, are influenced by anchoring, framing, representativeness, and egocentric bias. In short, arbitrators are like judges, and arbitral decisionmaking is like judicial decisionmaking. Whether appointed by the state and appearing in robes, or selected by parties and appearing in business suits, adjudicators are human beings, and human beings make predictable judgment and decisionmaking errors.

The insight that adjudicators, whether judges or arbitrators, will commit decision errors should inform those designing dispute systems, whether domestically or internationally. Those designing dispute resolution systems should focus less on who decides and more on structural features and procedural safeguards that increase the likelihood that the decisionmaker, whomever or whatever she is, provides justice.

---

<sup>262</sup> This observation may have limits, as experimental research suggests offering increased time does not enhance adjudication quality. Brian Sheppard, *Judging Under Pressure: A Behavioral Examination of the Relationship Between Legal Decisionmaking and Time*, 39 FLA. ST. U. L. REV. 931, 939 (2012).