

ORIGINAL ARTICLE

INTERNATIONAL LAW AND PRACTICE: SYMPOSIUM ON THE EFFECTS OF  
INTERNATIONAL JUDGES' PERSONAL CHARACTERISTICS ON THEIR JUDGING

## How social identity and social diversity affect judging

Lee Epstein<sup>1</sup> and Jack Knight<sup>2</sup>

<sup>1</sup>University of Southern California, USC Gould School of Law, 699 Exposition Boulevard, Los Angeles, CA 90089-0071, USA  
and <sup>2</sup>Duke University School of Law, 210 Science Drive, Durham, NC 27708, USA  
E-mails: [lepstein@law.usc.edu](mailto:lepstein@law.usc.edu), [knight@law.duke.edu](mailto:knight@law.duke.edu)

### Abstract

Judges like to claim that they are impartial decision-makers fully capable of suppressing their personal proclivities, as the rule of law requires. But a century's worth of studies undermines that view. Going under the name 'judicial behaviour', this vast literature shows that many extraneous (non-legal) factors affect the choices judges make. This article focuses on one strand of that literature – the effect of personal characteristics on judging, with emphasis on social identity and social diversity. We show that the literature is bifurcated: studies focusing on the *social identity of individual judges* (such as their gender, race, and nationality) generate findings consistent with in-group bias, whereas research on the *social diversity of judges sitting in panels* suggests that benefits can accrue from socially diverse courts. What the two sets of studies have in common, though, is just as important: both could make profound academic and policy contributions but require far more development if they are to realize their potential. We offer proposals for forward movement.

**Keywords:** collegial courts; gender; judges' social identity; judicial behaviour; nationality

### 1. Introduction

When commentators tout the value of the rule of law, they often emphasize the importance of 'impartiality' in judicial decision-making – that judges 'shall decide matters before them . . . on the basis of facts and in accordance with the law'<sup>1</sup> without regard to the identity of the parties.

For centuries this vision of impartiality has not only come to justify the rule of law and lend legitimacy to courts; it also lies at the core of most judges' self-image. Judges like to claim that they are dispassionate decision-makers fully capable of suppressing their 'personal proclivities'<sup>2</sup> And

<sup>1</sup>Basic Principles on the Independence of the Judiciary, GA/Res/40/146, UN Doc. A/RES/40/146 (13 December 1985), available at [www.un.org/ruleoflaw/blog/document/basic-principles-on-the-independence-of-the-judiciary/](http://www.un.org/ruleoflaw/blog/document/basic-principles-on-the-independence-of-the-judiciary/).

<sup>2</sup>E.g., A. Wistrich, J. Rachlinski and C. Guthrie, 'Heart Versus Head: Do Judges Follow the Law or Follow Their Feelings', (2015) 93 *Texas Law Journal* 855. In support of their assertion that 'most judges claim that they can effectively put emotions aside' (p. 860), the authors provide quotes from judges including US Supreme Court justice Antonin Scalia who wrote that 'good judges pride themselves on the rationality of their rulings and the suppression of their personal proclivities, including most especially their emotions'. A. Scalia and B. Garner, *Making Your Case: The Art of Persuading Judges* (2008), 32. The respected scholar and US appellate court judge Richard A. Posner concurred: 'most judges are (surprisingly to nonjudges) unmoved by the equities of the individual case': R. A. Posner, *How Judges Think* (2008), at 119.

yet a century's worth of studies undermines this claim.<sup>3</sup> Going under the name 'judicial behaviour', this vast literature shows that many extraneous (non-legal) factors affect the choices judges make.<sup>4</sup>

This article considers one strand of that literature, variously described as the analysis of the judges' social backgrounds,<sup>5</sup> personal attributes,<sup>6</sup> biographies,<sup>7</sup> or identities.<sup>8</sup> Whatever the descriptor, the basic idea is that personal characteristics affect judging.

The studies comprising this literature are both old and new. Early research tended to focus on career experience, asking, for example, whether former prosecutors-turned-judges are tougher on criminal defendants or whether judges who were corporate lawyers favour wealthy and powerful interests.<sup>9</sup> Partisan (political) identity too figured prominently in some of the original research, especially on US judges. Study after study found that judges appointed by, or affiliating with, the Democratic party, relative to Republican judges, were more supportive of workers, criminal defendants, and regulations on business – in other words, the Democratic judges were more left-leaning in their decisions.<sup>10</sup>

Emphasis on career paths and partisanship continues today.<sup>11</sup> At the same time, as judiciaries throughout the world have grown (somewhat) more diverse in composition,<sup>12</sup> the characteristics

<sup>3</sup>For a review of the early studies see L. Epstein, W. Landes and R. Posner, *The Behavior of Federal Judges* (2013), Ch. 2.

<sup>4</sup>These factors range from the judges' biographies to their ideological values to their quest to issue decisions that their society will respect. For a review see K. Weinshall and L. Epstein, *The Strategic Analysis of Judicial Behavior: A Comparative Perspective* (2021).

<sup>5</sup>J. Schmidhauser, 'Stare Decisis, Dissent, and the Backgrounds of Justices of the Supreme Court of the United States', (1962) 14 *University of Toronto Law Journal* 194; S. Ulmer, 'Social Background as an Indicator of the Votes of Supreme Court Justices in Criminal Cases: 1947–1956 Terms', (1973) 17 *American Journal of Political Science* 622; J. Brudney, S. Schiavoni and S. D. Merritt, 'Judicial Hostility Toward Labor Unions? Applying the Social Background Model to a Celebrated Concern', (1999) 60 *Ohio State Law Journal* 1675; T. George and T. Weaver, 'The Role of Personal Attributes and Social Backgrounds on Judging', in L. Epstein and S. Lindquist (eds.), *The Oxford Handbook of U.S. Judicial Behavior* (2017).

<sup>6</sup>George and Weaver, *ibid.*; C. Haines, 'General Observations on the Effects of Personal, Political, and Economic Influences in the Decisions of Judges', (1922) 17 *Illinois Law Review* 96; C. Tate, 'Personal Attribute Models of Voting Behavior of U.S. Supreme Court Justices: Liberalism in Civil Liberties and Economics Decisions, 1946–1978', (1981) 75 *American Political Science Review* 355.

<sup>7</sup>See generally R. Posner, 'Judicial Biography', (1995) 70 *New York University Law Review* 502.

<sup>8</sup>Epstein and Weinshall, *supra* note 4.

<sup>9</sup>E.g., C. Tate and P. Sittiwong, 'Decision Making in the Canadian Supreme Court: Extending the Personal Attributes Model Across Nations', (1989) 51 *Journal of Politics* 900. For a list of studies on career experience see L. Epstein, J. Knight and A. Martin, 'The Norm of Prior Judicial Experience and Its Consequences for Career Diversity on the U.S. Supreme Court', (2003) 91 *California Law Review* 903, 961–5.

<sup>10</sup>Early studies include G. Schubert, *Quantitative Analysis of Judicial Behavior* (1959); S. Nagel, 'Political Party Affiliation and Judges' Decisions', (1961) 55 *American Political Science Review* 843; S. Ulmer, 'The Political Party Variable on the Michigan Supreme Court', (1962) 11 *Journal of Public Law* 352.

<sup>11</sup>E.g., J. Stribopoulos and M. Yahya, 'Does a Judge's Party of Appointment or Gender Matter to Case Outcomes: An Empirical Study of the Court of Appeal for Ontario', (2007) 45 *Osgoode Hall Law Journal* 315; Epstein, Landes and Posner, *supra* note 3; A. Cohen and C. Yang, 'Judicial Politics and Sentencing Decisions', (2019) 11 *American Economic Journal: Economic Policy* 160; C. Liu, 'Judge Political Party Affiliation and Impacts on Corporate Environmental Litigation', (2020) 64 *Journal of Corporate Finance* 101670.

<sup>12</sup>M. Williams and F. Thames, 'Women's Representation on High Courts in Advanced Industrialized Countries', (2008) 4 *Politics & Gender* 451; S. Basabe-Serrano, 'The Representation of Women in the Judicial Branch: Eighteen Latin American High Courts in Comparative Perspective', (2019) 185 *Revista de Estudios Políticos* 259. See also *infra* Section 2.

under analysis have expanded to include the judges' (and, sometimes, the litigants') social categories<sup>13</sup> – especially gender,<sup>14</sup> religion,<sup>15</sup> race,<sup>16</sup> ethnicity,<sup>17</sup> and nationality<sup>18</sup>

Increasing attention to social identity and social diversity has contributed to the study of judging in more ways than one. First, because of world-wide interest in these twin topics, the field is more interdisciplinary than ever. In the not-so-distant past, the most prominent studies were authored by US political scientists studying US courts.<sup>19</sup> But these days work is just as likely to draw on economics,<sup>20</sup> law,<sup>21</sup> psychology,<sup>22</sup> and sociology,<sup>23</sup> as it is on political science; and research is increasingly global in scope, covering judges serving on domestic and international courts.<sup>24</sup> As diversity in approach, tools, and targets of inquiry has increased, so too has the depth of the studies and thus their capacity to enrich explanations of judging.<sup>25</sup>

A second contribution traces to the studies themselves. Reading across the growing number of articles and books on social identity and social diversity in courts, it seems that their findings are converging – though in distinct ways. On the one hand, research that characterizes *individual judges* on the basis of their *social identity* (gender, race, nationality, and so on) tends to generate results in line with in-group bias: the tendency of individuals to favour members of their own group over outsiders (such as, international court judges' favouritism toward their home government<sup>26</sup>). On the other hand, research on the *social diversity* of *collegial courts* (those on which judges sit in panels or *en banc*) suggests that greater heterogeneity can produce benefits in the form of better decisions.

Characterized in this way, the two lines of research have quite different, even conflicting, implications for the legitimacy of courts in particular and the rule of law in general. At the level of the

<sup>13</sup>See generally A. Harris and M. Sen, 'Bias and Judging', (2019) 22 *Annual Review of Political Science* 241.

<sup>14</sup>E.g., C. Boyd, L. Epstein and A. Martin, 'Untangling the Causal Effects of Sex on Judging', (2010) 54 *American Journal of Political Science* 389; C. Boyd, 'Representation on the Courts? The Effects of Trial Judges; Sex and Race', (2016) 69 *Political Research Quarterly* 788; E. Voeten, 'Gender and Judging: Evidence from the European Court of Human Rights', (2021) 28 *Journal of European Public Policy* 1453.

<sup>15</sup>E.g., D. Koev, 'Not Taking it on Faith: State and Religious Influences on European Court of Human Rights Judges in Freedom of Religion Cases', (2019) 18 *Journal of Human Rights* 184; M. Shayo and A. Zussman, 'Judicial Ingroup Bias in the Shadow of Terrorism', (2011) 126 *Quarterly Journal of Economics* 1447; O. Gazal-Ayal and R. Sulitzeanu-Kenan, "'Let My People Go": Ethnic In-Group Bias in Judicial Decisions—Evidence from a Randomized Natural Experiment', (2010) 7 *Journal of Empirical Legal Studies* 403.

<sup>16</sup>M. Sen, 'Is Justice Really Blind? Race and Appellate Review in U.S. Courts', (2015) 44 *Journal of Legal Studies* 187; Boyd, Epstein and Martin *supra* note 14; J. Rachlinski et al., 'Does Unconscious Racial Bias Affect Trial Judges?', (2009) 84 *Notre Dame Law Review* 1195.

<sup>17</sup>Shayo and Zussman, *supra* note 15; Gazal-Ayal and Sulitzeanu-Kenan, *supra* note 15; A. Schwartz and M. Murchison, 'Judicial Impartiality and Independence in Divided Societies: An Empirical Analysis of the Constitutional Court of Bosnia-Herzegovina', (2016) 51 *Law and Society Review* 821.

<sup>18</sup>E. Posner and M. de Figueiredo, 'Is the International Court of Justice Biased?', (2005) 34 *Journal of Legal Studies* 599; E. Voeten, 'The Impartiality of International Judges: Evidence from the European Court of Human Rights', (2008) 102 *American Political Science Review* 417; Koev, *supra* note 15; Schwartz and Murchison, *supra* note 17. Worth noting is that the analysis of the effect of national identity on international courts' decisions has a long pedigree, in part because those courts have always been diverse nationality-wise. Examples include T. Hensley, 'National Bias and the International Court of Justice', (1968) 12 *Midwest Journal of Political Science* 568; M. Kuijer, 'Voting Behavior and National Bias in the European Court of Human Rights and the International Court of Justice', (1997) 10 *Leiden Journal of International Law* 49. See also *infra* Section 2.1.

<sup>19</sup>See, e.g., studies listed in *supra* notes 5, 9, 10.

<sup>20</sup>E.g., Posner and de Figueiredo, *supra* note 18. For a review, Epstein and Weinsahl, *supra* note 4.

<sup>21</sup>E.g., A. Cox and T. Miles, 'Judging the Voting Rights Act', (2008) 108 *Columbia Law Review* 1.

<sup>22</sup>E.g., Rachlinski et al., *supra* note 16.

<sup>23</sup>E.g., R. Gill, M. Kagan and F. Marouf, 'The Impact of Maleness on Judicial Decision Making: Masculinity, Chivalry, and Immigration Appeals', (2017) 7 *Politics, Groups, and Identities* 509.

<sup>24</sup>See studies in *supra* note 18.

<sup>25</sup>See generally L. Epstein, U. Šadl and K. Weinsahl, 'The Role of Comparative Law in the Analysis of Judicial Behavior', *American Journal of Comparative Law*, forthcoming.

<sup>26</sup>See *supra* note 18.

*individual judge*, differences in decision making based on *social identity* challenge an underlying justification for the rule of law: that it fosters impartiality. At the *collective level of collegial courts*, *social diversity* may enhance the quality of judicial decisions, thereby generating greater respect for judges and increasing the legitimacy of their courts.

These are the differences between studies on social identity and on social diversity. What the two have in common, though, may be even more important: Both have the potential to make important contributions to public policy.

Starting with work on the individual judge: if research can isolate features of judging that are most likely to exacerbate differences emerging from social identity, then scholars should be able to propose institutional reforms that will diminish their significance. An obvious example focuses on the connections among institutions (formal and informal rules) governing the appointment of judges, the people selected to serve and ultimately, the choices they, as judges, make. Understanding how the judges' personal characteristics affect their choices could inform institutional design by, say, identifying selection procedures that reduce the effect of partisan political identity on the judges' decisions or that lead to greater gender diversification.<sup>27</sup> Along somewhat different lines, research on social identity flagging particular biases – such as the tendency of international court judges to favour their home government when it is a party to a dispute<sup>28</sup> – ought to force judges to confront the difficulties in suppressing 'personal proclivities' and emotions.<sup>29</sup>

Turning to *collegial courts*, research could determine whether greater social diversity, in fact, has salutary effects and what institutional practices are most likely to generate those effects. The basic task here would be to use empirical research to craft institutional procedures that foster positive collective benefits.<sup>30</sup>

However conceivably important these policy implications – whether for the social identity of individual judges or the social diversity of collegial courts – 'conceivably' is the operative word: the implications remain more aspirational than realized. For even with the growth of research on the judges' personal characteristics, much more work is needed for the field to reach its full potential in terms of policy impact and we might add, academic consequence. For this reason, we propose opportunities for forward movement.

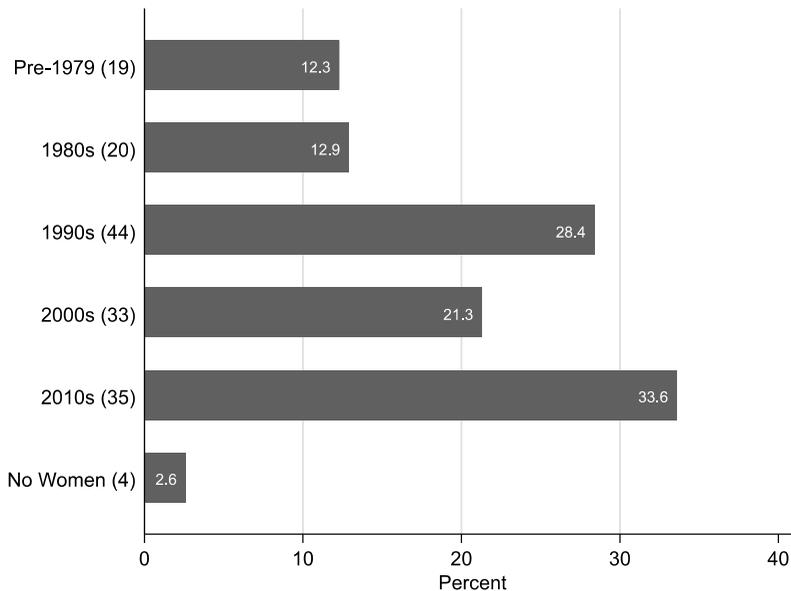
The discussion unfolds in two parts, reflecting the bifurcated approach (and implications) of existing studies of identity, on the one hand, and diversity, on the other. Section 2 focuses on identity, detailing research that characterizes individual judges as members of a particular group based on their gender, race, nationality, and so on. That literature, as we just suggested, generates findings consistent with in-group bias. Section 3, on diversity, considers studies of collegial courts. That work mostly reaches results consistent with the benefits that can accrue in socially diverse teams or groups, though again the research is quite nascent. Along the way, we point out gaps in

<sup>27</sup>Some of this work has begun in earnest. On partisanship see B. Alarie and A. J. Green, *Commitment and Cooperation on High Courts: A Cross-Country Examination of Institutional Constraints on Judges* (2017) (removing political actors from the selection process leads to less partisan or ideological courts). On gender, see Williams and Thames, *supra* note 12; M. Valdini and C. Shortell, 'Women's Representation in the Highest Court: A Comparative Analysis of the Appointment of Female Justices', (2016) 69 *Political Research Quarterly* 865; I. Araya, M. Hughes and A. Pérez-Liñán, 'Judicial Reshuffles and Women Justices in Latin America', (2021) 65 *American Journal of Political Science* 373; M. Escobar-Lemmon et al. 'Breaking the Judicial Glass Ceiling: The Appointment of Women to High Courts Worldwide', (2021) 83 *Journal of Politics* 662; N. Arrington et al., 'Constitutional Reform and the Gender Diversification of Peak Courts', (2021) 115 *American Political Science Review* 851.

<sup>28</sup>See the studies in *supra* note 18.

<sup>29</sup>See, e.g., Wistrich, Rachlinski and Guthrie, *supra* note 2, at 909 ('Judges should be cognizant of their [reliance on the affect heuristic]. Most people fail to recognize its hidden influence. Awareness is not sufficient to ensure that judges keep emotional responses in check, but it is a necessary first step.')

<sup>30</sup>Running along these lines are studies on 'panel effects' such as S. Haire, L. Moyer and S. Treier, 'Diversity, Deliberation, and Judicial Opinion Writing', (2013) 1 *Journal of Law and Courts* 303; R. Hunter, 'More than Just a Difference Face? Judicial Diversity and Decision-making', (2015) 68 *Current Legal Problems* 119. See also *infra* Section 3.1.



**Figure 1.** When the first woman was selected to serve on the highest court in 155 countries. The number in parentheses is the number of countries. E.g., before 1979, in only 19 of the 155 countries (12.3 per cent) did the first woman serve; in the 1980s, 20 countries (12.9 per cent) selected their first woman.

the literature and offer suggestions for scholars hoping to contribute to these exciting areas of study.

## 2. Social identity and the judge

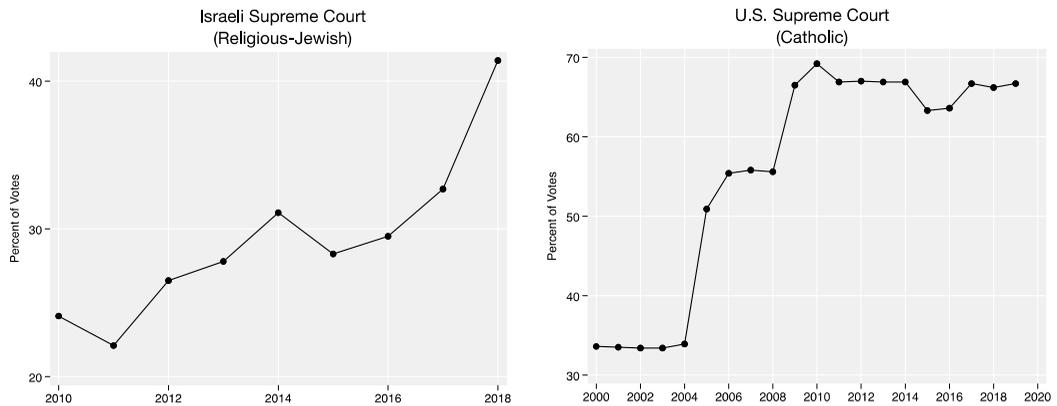
Prior to the 1980s, the systematic study of social identity was difficult because courts in many countries were quite homogeneous. Consider, for example, the judges' gender. As Figure 1 shows,<sup>31</sup> through the 1970s only about 12 per cent of the 155 countries represented in the data had selected a woman – their first woman – to sit on the country's highest court; in the 1980s, that percentage doubled, such that at least one woman had served on about  $(12.3 + 12.9 =)$  25 per cent of the countries' courts. By the end of 2010s, only four countries had never appointed a woman to their highest court.

This is not to say that all high courts are now composed of substantial percentages of female judges. On the contrary: the fraction varies considerably. For international courts, Grossman's (1999–2015) data show that women occupied 47 per cent of the seats on the International Court of Justice but only 14 per cent on the European Court of Justice.<sup>32</sup> Valdini and Shortell's (2017) data for domestic high courts also shows considerable variation, from Latvia's 57.1 per cent female to under 10 per cent for 11 countries (including Italy, Panama, India, and the UK).<sup>33</sup> Nonetheless, even low percentages can mask growth in gender diversity over time.

<sup>31</sup>Data are from Arrington et al., *supra* note 27.

<sup>32</sup>N. Grossman, 'Achieving Sex-Representative International Court Benches', (2016) 110 *American Journal of International Law* 82, at 83. Grossman attributes the variation, in part, to legal requirements mandating the consideration of gender when selecting judges: 'Of the five courts with the highest percentage of women on the bench from 1999 to 2015, four had either aspirational statements for inclusion or quotas' (at 93). For other explanations of gender variation see the studies listed in *supra* note 27.

<sup>33</sup>Valdini and Shortell, *supra* note 27.



**Figure 2.** Percentage of votes cast by religious-Jewish justices on the Israeli Supreme Court (left panel) and by Catholic justices on the US Supreme Court (right panel).

In Latin America, for example, the number of female justices ‘increased dramatically . . . from 3% of all justices in the region’s high courts in 1980 to 19% in 2010’.<sup>34</sup>

Quantifying increases in other social categories (such as race, ethnicity, and religion) is harder because of varying cleavages on the relevant identities, as well as a lack of data on judge characteristics.<sup>35</sup> But a look at religion in Israel and the United States may suffice to make the point. As Figure 2 (left panel) shows,<sup>36</sup> in 2010 religious Jewish justices (relative to secular Jewish and non-Jewish justices) cast fewer than 25 per cent of the votes in the Israeli Supreme Court; a decade later, the percentage jumped to 41. Growth in religious diversity on the US Supreme Court is equally apparent (right panel). Between 2000 and 2020, the percentage of votes cast by Catholic justices (relative to all other religions) doubled, from about 33 to over 66.

## 2.1 The studies

Increasing social diversity among judges – not just on gender and religion but also on race, nationality, and ethnicity – naturally enough generated research centred on whether the particular category ‘mattered’. Do female judges ‘speak in a different voice’ than their male counterparts?<sup>37</sup> Are Black US judges inclined to treat Black defendants more leniently (and likewise for white judges/defendants)?<sup>38</sup> Compared to less devout judges, do religious judges more frequently support claims of religious liberty?<sup>39</sup> The diversity turn also triggered new research programmes on courts

<sup>34</sup>Araya, Hughes and Pérez-Liñán, *supra* note 27, at 373.

<sup>35</sup>This may change with the release of Judicial BioStats ([www.judicialbiostats.org](http://www.judicialbiostats.org)) – a project that is compiling biographical data on judges worldwide.

<sup>36</sup>Data on the Israeli justices developed by Weinshall and Epstein, *supra* note 4; data on the US justices is from L. Epstein and E. Posner, ‘The Roberts Court and the Transformation of Constitutional Protections for Religion: A Statistical Portrait’, (2022) 2021 *Supreme Court Review* 315.

<sup>37</sup>The ‘different voice’ language is from C. Gilligan, *In a Different Voice: Psychological Theory and Women’s Development* (1982). In the judging-gender literature, the expectation from Gilligan’s work seems to be that female judges will bring a ‘feminine perspective’ to the bench – one that ‘encompasses all aspects of society, whether or not they affect men and women differently’ and not only ‘the political agenda associated with feminism’. S. Sherry, ‘Civic Virtue and the Feminine Voice in Constitutional Adjudication’, (1986) 72 *Virginia Law Review* 543. As discussed below, empirical studies do not seem to support the hypothesis of differences between male and female judges across all issues.

<sup>38</sup>E.g., S. Welch, M. Combs and J. Gruhl, ‘Do Black Judges Make a Difference?’, (1988) 32 *American Journal of Political Science* 126.

<sup>39</sup>E.g., K. Weinshall, ‘Courts and Diversity: Normative Justifications and their Empirical Implications’, *Law & Ethics of Human Rights* (forthcoming).

that always have been socially diverse on some dimension – on the European Court of Human Rights, for example, nearly 50 different nationalities have been represented at one time or another.<sup>40</sup> But studies *systematically* exploring whether national identity (or ethno-national affiliation) affects the choices of judges serving on international and domestic courts have grown in number and sophistication in recent years.<sup>41</sup>

Theoretically, projects attempting to answer questions about (possible) differences based on social categories are all over the map, offering an array of explanations.<sup>42</sup> The findings, though, are more uniform, with many fitting comfortably with one of the most central and best-documented manifestations of social group identity: in-group bias (or favouritism). More than five decades ago, social scientists noticed the tendency of individuals to favour members of their own group over outsiders;<sup>43</sup> and now there is a very substantial literature on the subject across law and the (social) sciences.<sup>44</sup> The result is an impressive body of evidence showing that people tend to be more helpful, more willing to allocate resources, and more supportive of policies advocated by members of their own group (and vice versa for the outgroup). Moreover, the greater the salience and relevance of the grouping to the task at hand, ‘the stronger these intergroup divisions’.<sup>45</sup>

Evidence consistent with in-group bias in courts – the ‘us-against-them’ judging at the heart of the problematic relationship between impartiality and judicial decisions – is far and wide.<sup>46</sup> Starting with gender, contemporary studies have converged on results showing that female judges are more favourable toward plaintiffs (mostly women) in cases of gender-based discrimination and sexual harassment. This holds on courts as diverse as the US intermediate appellate courts,<sup>47</sup> the European Court of Human Rights,<sup>48</sup> and the Supreme Court of Canada.<sup>49</sup> Then there is evidence in the family law context of female judges tending to favour mothers over fathers, whereas male judges support fathers over mothers;<sup>50</sup> likewise, in criminal cases involving sex offences (again,

<sup>40</sup>Data supplied by Eric Voeten at Georgetown University.

<sup>41</sup>See the studies listed in *supra* note 18.

<sup>42</sup>The accounts range from the representational (judges serve as representatives of their social group and so work to protect it in litigation of direct interest) to the informational (judges of different social groups possess unique and valuable information emanating from shared experiences) to the organizational (judges of all social groups in a society undergo identical professional training, obtain their jobs through the same procedures, and confront similar constraints once on the bench). See Boyd, Epstein and Martin, *supra* note 14.

<sup>43</sup>H. Tajfel, ‘Experiments in Intergroup Discrimination’, (1970) 223 *Scientific American* 96; M. Billig and H. Tajfel, ‘Social Categorization and Similarity in Intergroup Behavior’, (1973) 3 *European Journal of Social Psychology* 27; H. Tajfel and J. Turner, ‘An Integrative Theory of Intergroup Conflict’, in W. Austin and S. Worchel (eds.), *The Social Psychology of Intergroup Relations* (1979), at 33.

<sup>44</sup>For a general review see M. Hewstone, M. Rubin and H. Willis, ‘Intergroup Bias’, (2002) 53 *Annual Review of Psychology* 575; for a review with applications to judging see L. Epstein, C. Parker and J. Segal, ‘Do Justices Defend the Speech They Hate?: An Analysis of In-Group Bias on the U.S. Supreme Court’, (2018) 6 *Journal of Law and Courts* 237.

<sup>45</sup>S. Iyengar et al., ‘The Origins and Consequences of Affective Polarization in the United States’, (2019) 22 *Annual Review of Political Science* 129, at 130.

<sup>46</sup>There are exceptions. For example, in a study of decisions in immigration cases, Gill, Kagan and Marouf, *supra* note 23, find effects seemingly at odds with in-group bias. As the authors explain, ‘Masculinity theory predicts that men will be disadvantaged when appearing in front of an all-male panel, while chivalry theory predicts that women will benefit from an all-male panel. Our results provide support for both of these theories.’

<sup>47</sup>Boyd, Epstein and Martin, *supra* note 14; L. Moyer and H. Tankersley, ‘Judicial Innovation and Sexual Harassment Doctrine in the US Courts of Appeals’, (2012) 65 *Political Research Quarterly* 784; J. Peresie, ‘Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts’, (2005) 114 *Yale Law Journal* 1759. Then again, see A. Glynn and M. Sen, ‘Identifying Judicial Empathy: Does Having Daughters Cause Judges to Rule of Women’s Issues’, (2015) 59 *American Journal of Political Science* 37. The authors demonstrate that male US appellate judges with daughters decide cases implicating gender in a ‘more feminist fashion’ than male judges who have only sons.

<sup>48</sup>Voeten, *supra* note 14.

<sup>49</sup>S. Johnson and D. Songer, ‘Judge Gender and the Voting Behavior of Justices on Two North American Supreme Courts’, (2009) 30 *Justice System Journal* 265.

<sup>50</sup>E.g., Cohen and Yang, *supra* note 11.

with mostly male defendants), female judges favour the state at greater rates than males.<sup>51</sup> Beyond these types of cases – *all salient to gender* – few (or weaker) differences emerge between male and female judges. Running along similar lines are studies of the judges' (and litigants') race in the United States. To supply but a few examples: White judges are less likely than Black judges to hold for Black defendants claiming police misconduct;<sup>52</sup> Black federal appeals judges, relative to whites, tend to support plaintiffs of colour in voting rights cases<sup>53</sup> and Black claimants in employment discrimination litigation;<sup>54</sup> and white judges are less supportive of affirmative action (diversity) programs than Black judges.<sup>55</sup> Worthy of mention too is Sen's research showing a greater willingness on the part of (mostly white) US appellate judges to reverse the decisions of Black trial court judges.<sup>56</sup>

Studies of religion and nationality also shore up various forms of us-against-them judging. In separate papers, Weinshall (Israeli justices), and Epstein and Posner (US justices) find a relationship between devoutness and support for religion.<sup>57</sup> Religious Jewish justices on the Israeli Court are far less likely to overrule decisions of the rabbinical court than secular Jews; and highly devout US justices (many of whom are Catholic) more frequently rule in favour of religious liberty than the less devout.<sup>58</sup> Koev's study of religious freedom cases in the European Court of Human Rights also points to a form of us-against-them judging: the judges are less likely to rule in favour of Muslim applicants.<sup>59</sup> Whether this result reflects bias against adherents of Islam is hard to know, as Koev acknowledges. Nonetheless, 'the strength of the empirical relationship, when paired with the salient and controversial role Muslims (especially Muslim migrants) play in Europe's contemporary political climate, suggests such a bias is plausible'.<sup>60</sup>

In-group favouritism based on national identity is equally evident. Empirical studies give little reason to doubt 'that the nationality of the international adjudicator matters'.<sup>61</sup> To provide an (extreme) example, an Italian judge on the European Court of Human Rights dissented in 133 judgments concerning alleged Italian violations of the European Convention on Human Rights.<sup>62</sup> Even on domestic courts (ethno-)national affiliation can lead to in-group bias, as a study of the Constitutional Court of Bosnia-Herzegovina uncovered: 'the judges . . . divide predictably along ethno-national lines [and] these divisions cannot be reduced to a residual loyalty to their appointing political parties'.<sup>63</sup>

<sup>51</sup>Weinshall, *supra* note 39, for example, finds that female judges on the Israeli Supreme Court, 'were significantly "tougher on crime" when sex offenses were involved, ruling for the state in 80.2% of their decisions, compared to 67.1% of the decisions by men'.

<sup>52</sup>N. Scherer, 'Blacks on the Bench', (2004) 119 *Political Science Quarterly* 655.

<sup>53</sup>Cox and Miles, *supra* note 21.

<sup>54</sup>J. Morin, 'The Voting Behavior of Minority Judges in the US Courts of Appeals: Does the Race of the Claimant Matter?', (2014) 43 *American Politics Research* 34.

<sup>55</sup>J. Kastellec, 'Racial Diversity and Judicial Influence on Appellate Courts', (2013) 57 *American Journal of Political Science* 167.

<sup>56</sup>Sen, *supra* note 16. Sen notes that 'Although having blacks on the reviewing panel appears to attenuate the effect, there are too few black appeals court judges to make meaningful inferences.'

<sup>57</sup>Epstein and Posner, *supra* note 36; Weinshall, *supra* note 39.

<sup>58</sup>See also W. Blake, 'God Save This Honorable Court: Religion as a Source of Judicial Policy Preferences', (2012) 65 *Political Research Quarterly* 814 (Catholic justices form a distinctive voting bloc, voting consistently with Catholic theological commitments); L. Wasserman and J. Hardy, 'U.S. Supreme Court Justices' Religious and Party Affiliation', (2013) 2 *British Journal of America Legal Studies* 111 (Catholic justices favour religious organizations more than Protestant justices – by a factor of 12).

<sup>59</sup>Koev, *supra* note 15.

<sup>60</sup>*Ibid.*, at 197.

<sup>61</sup>C. Titi, 'Nationality and Representation in the Composition of the International Bench: Lessons from the Practice of International Courts and Tribunals and Policy Options for the Multilateral Investment Court', *SSRN*, 29 January 2020, available at [ssrn.com/abstract=3519863](https://ssrn.com/abstract=3519863). Titi references some of the studies listed in *supra* note 18.

<sup>62</sup>*Ibid.*

<sup>63</sup>Schwartz and Murchison, *supra* note 17.

## 2.2 The gaps

However promising this line of research, much more work must be done to develop a fuller picture of the relationship between social identity and judging. Four gaps come to mind.

First, although many results are *consistent* with in-group bias accounts, ‘consistent’ is worthy of emphasis. With very few exceptions,<sup>64</sup> the researchers did not set out directly to assess us-against-them judging. They instead sought to answer questions that focus on difference (e.g., do female judges speak in a different voice?), and not on the judges’ treatment of out- versus in- groups.

The exceptions, though, are instructive. Take Shayo and Zussman’s famous study of small claims courts in Israel, where cases are randomly assigned to the judges.<sup>65</sup> The analysis unearths clear ‘us-against-them’ judging based on religion and ethnicity: Jewish judges systematically favour Jewish litigants, and Arab judges favour Arab litigants. Further, when the salience of these identities increases, say, in the aftermath of terrorist attacks, the bias strengthens.

Because of its careful attention to theory, design, and analysis – not to mention the researchers’ substantive knowledge – this study provides a clear roadmap for future work on judicial in-group bias.<sup>66</sup> Indeed, to the extent that in-group bias is a feature of human cognition, Shayo and Zussman’s findings should transport to other societies, though the relevant ‘teams’ may differ. In the US, for example, where the American public is highly polarized along partisan lines, one might expect judges to be biased in favour of co-partisans and against opposing partisans in, say, disputes implicating elections and voting.

Second and relatedly, assuming judges are no more immune to in-group bias than the rest of us, the mechanism remains unclear. The chief hurdle is that in-group bias could be a manifestation of judges rationally pursuing their self-interest, for example, voting for people like them to advance their future job prospects; or of judges ‘thinking fast’,<sup>67</sup> that is, relying on heuristics, emotions, intuitions, and the like to make quick decisions without much effort. In other words, looming large is the problem of behavioural (or observational) equivalence – when different accounts generate the same prediction about the judges’ behaviour.

By way of example, consider Posner and de Figueiredo’s study on the International Court of Justice (ICJ).<sup>68</sup> After demonstrating that ICJ judges tend to vote in favour of their home country (a form of in-group bias), the authors offer an explanation grounded in (self-interested) rational choice theory:

Economically, judges may be motivated by material incentives. Judges who defy the wills of their government by holding against it may be penalized. The government may refuse to support them for reappointment and also refuse to give them any other desirable government position after the expiration of their term.<sup>69</sup>

But another mechanism is equally plausible: ICJ judges side with their own country not because they are rationally advancing an economic or any other interest but because of an emotional response. Posner and de Figueiredo recognize as much when they offer this alternative explanation for their finding:

<sup>64</sup>Wistrich, Rachlinski and Guthrie, *supra* note 29; Epstein, Parker and Segal, *supra* note 44; Shayo and Zussman, *supra* note 15.

<sup>65</sup>Shayo and Zussman, *ibid.*

<sup>66</sup>See also M. Shayo, ‘A Model of Social Identity with an Application to Political Economy: Nation, Class and Redistribution’, (2009) 103 *American Political Science Review* 147 (offering a theoretical approach for modeling the activation of social identities).

<sup>67</sup>See generally D. Kahneman, *Thinking, Fast and Slow* (2011).

<sup>68</sup>Posner and de Figueiredo, *supra* note 18.

<sup>69</sup>*Ibid.*, at 608.

Psychologically, if judges identify with their countries, they may find it difficult to maintain impartiality. ICJ judges are not only nationals who would normally have strong emotional ties with their country; they also have spent their careers in national service as diplomats, legal advisors, administrators, and politicians. Even with the best intentions, they may have trouble seeing the dispute from the perspective of any country but that of their native land.<sup>70</sup>

Differentiating between these two possible mechanisms is not easy because either could generate the observed results: in-group favoritism. With clever designs and data, however, we can and, more to the point, must solve the equivalence problem if we are to develop more comprehensive explanations of the role of social identity in judging.

A third gap in the social identity literature squarely implicates ‘identity’ or more accurately ‘identities’. Virtually all existing data studies of judging ‘separate out the effects of individual identity attributes’.<sup>71</sup> A Jewish female judge is decomposed into a Jewish effect and a female effect; a Black male British judge into a Black, a male, and a British effect, and on and on. Many scholars working in the diversity space, however, reject this approach arguing instead that ‘the inseparability of attributes . . . should give us pause when interpreting data sorted by a single attribute’.<sup>72</sup>

Recognizing that individual judges are, like all of us, bundles of identities – identities that intersect and overlap<sup>73</sup> – is crucial to advance work in the field. But the studies need not move in lock-step. One approach, following from research on social categories (with Shayo’s theoretical work especially informative<sup>74</sup>), is to consider more carefully when one identity over others is likely to be activated because it may be particularly salient to the dispute at hand. Another is explicitly to account for the intersectionality of identities in the research.<sup>75</sup> One of the rare examples is a study of US appellate judges that models the joint effects of gender and race.<sup>76</sup>

A final possibility for future work entails a synthesis of research on judicial selection mechanisms and research on social identity. Existing studies on the effects of selection mechanisms have focused on the general question of whether different mechanisms produce differences in decision making. The studies have yet to drill down to investigate the substantive content of those differences. Future work on the institutional design of judicial selection should contemplate whether different methods of appointment (and retention) result in judges who are more or less prone to these forms of in-group bias.

<sup>70</sup>*Ibid.* See also Kuijer, *supra* note 18, at 49, quoting Report of the Committee on the revision of the Rules of Court, PCIJ Rep., Series E No.4, 75 (1927) (‘Of all influences to which men are subject, none is more powerful, more pervasive or more subtle than the tie of allegiance that binds them [judges] to the land of their homes and kindred and to the great sources of the honors and preferments for which they are so ready to spend their fortunes and to risk their lives.’).

<sup>71</sup>S. Page, *The Diversity Bonus* (2017), 142. E. Bonthuys, ‘Gender and Race in South African Judicial Appointments’, (2015) 23 *Feminist Legal Studies* 127, makes a similar point.

<sup>72</sup>Page, *ibid.*, at 142.

<sup>73</sup>K. Crenshaw coined the term ‘intersectionality’ in her ‘Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics’, (1989) *University of Chicago Legal Forum* 139. And the term has now made its way into the *Oxford English Dictionary*, which defines intersectionality as ‘The interconnected nature of social categorizations such as race, class, and gender, regarded as creating overlapping and interdependent systems of discrimination or disadvantage of social identities.’ Studies exploring intersectionality are many in number; indeed, Crenshaw’s article has been cited more than 23,000 times (according to a Google Scholar search run in December 2021). For an introduction to the topic see P. Collins and S. Bilge, *Intersectionality* (2020). Highly cited studies include L. McCall, ‘The Complexity of Intersectionality’, (2005) 30 *Signs* 771; N. Davis, ‘Intersectionality and Feminist Politics’, (2006) 13 *European Journal of Women’s Studies* 193; S. Cho, K. Crenshaw and L. McCall, ‘Toward a Field of Intersectionality Studies: Theory, Applications, and Praxis’, (2013) 38 *Signs* 785.

<sup>74</sup>Shayo, *supra* note 66.

<sup>75</sup>See A. Kang et al., ‘Diverse and Inclusive High Courts: A Global and Intersectional Perspective’, (2020) 8 *Politics, Groups, & Identities* 812 (analysing attempts at ‘intersectional inclusion’ in the high courts of Canada and South Africa).

<sup>76</sup>T. Collins and L. Moyer, ‘Gender, Race, and Intersectionality on the Federal Appellate Bench’, (2008) 61 *Political Research Quarterly* 219. See also S. Haire and L. Moyer, *Diversity Matters: Judicial Policy Making in the U.S. Courts of Appeals* (2015).

### 3. Social diversity on collegial courts

Depending on your perspective, judicial in-group bias could be seen as offsetting years of ‘exclusionary legal processes’<sup>77</sup> or it could be seen as violating a guiding principle of most courts: to treat all parties equally. Either way, a mix of social identities is essential for ensuring socially diverse courts (‘teams of judges’), which, in turn, may be crucial for developing innovative, high-quality solutions to the kinds of complex problems that confront contemporary courts. In other words, a ‘diversity bonus’<sup>78</sup> may follow for courts with a mix of social identities.<sup>79</sup>

Why? Diversity theorists posit several mechanisms. Most well suited to courts, we believe, are approaches that emphasize shared information and the requirement of reaching an agreement on a collective decision. Scott E. Page, for example, argues that the key to building great teams lies in germane cognitive diversity, which amounts to ‘differences in information, knowledge, representations, mental models, and heuristics’ that team members bring to the tasks of ‘problem-solving, predicting, and innovating’.<sup>80</sup> To the extent that social category diversity – differences in race, gender, ethnicity, religion, nationality – feeds into people’s ‘cognitive repertoire’, socially diverse groups will perform better.

Another prominent scholar, Katherine W. Phillips, agrees with Page but more directly connects social diversity and ‘smarter’ decision making. As she puts it:

The key to understanding the positive influence of diversity is the concept of informational diversity. When people are brought together to solve problems in groups, they bring different information, opinions and perspectives. This makes obvious sense when we talk about diversity of disciplinary backgrounds—think [of an] interdisciplinary team building a car. The same logic applies to social diversity. People who are different from one another in race, gender and other dimensions bring unique information and experiences to bear on the task at hand. A male and a female engineer might have perspectives as different from one another as an engineer and a physicist—and that is a good thing.<sup>81</sup>

On this account, social diversity leads to better decisions as people bring different perspectives to bear on the problem at hand; in other words, the more diverse the inputs, the stronger the outputs. Benefits also accrue, Phillips maintains, ‘from simply adding social diversity to a group’ because ‘people believe that differences of perspective might exist among them and that belief makes people change their behavior’ by working harder and ‘encouraging the consideration of alternatives before any interpersonal interaction takes place’.<sup>82</sup>

The mechanisms identified by Page and Phillips are especially relevant to analyses of collegial courts. Diverse judges likely bring different experiences and perspectives to the factual and legal questions they confront, and the potential influences of those differences are numerous. They may affect the ways in which a collegial court identifies and weighs the controlling facts in a particular case. They may influence how the court assesses the implications of the relevant ‘law’ (broadly defined to include constitutional provisions, statutes, past judicial decisions, and the like) for the kinds of subsequent behaviour its decisions might induce. Or they might shape the court’s collective assessment of what a fair or just outcome entails.

<sup>77</sup>Escobar-Lemmon et al., *supra* note 27.

<sup>78</sup>Page, *supra* note 71. See also Page’s *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies* (2007). Section 3.2 reviews empirical studies supporting this claim – and how the lessons from those studies could be applied to courts.

<sup>79</sup>Almost needless to write, many justifications exist for diversity, inclusion, and equity. In the context of courts see, e.g., N. Grossman, ‘Sex on the Bench: Do Women Judges Matter to the Legitimacy of International Courts’, (2011) 12 *Chicago Journal of International Law* 647; Haire and Moyer, *supra* note 76; Epstein, Martin and Knight, *supra* note 9.

<sup>80</sup>Page, *supra* note 71, at 14–15.

<sup>81</sup>K. Phillips, ‘How Diversity Makes Us Smarter’, (2014) 311 *Scientific America* 42.

<sup>82</sup>*Ibid.*

These are just a few of the potential effects of diversity on the collective decision-making process of a collegial court. Whether or not they, in fact, enhance collective decisions in the beneficial ways anticipated by the ‘diversity bonus’ literature is an important subject for empirical research.

### 3.1 The studies

Although research directly putting ideas about the value of diversity in the courtroom to the test is scant, a handful of results could be seen as consistent with them. These results follow from ‘The Collegial Court’ studies – a line of inquiry that seeks to assess whether a case’s outcome (or a judge’s vote) would have been different had a single judge, and not a panel, decided the case. In other words, the key question is whether ‘collegial’ (or ‘panel’ or ‘peer’) effects exist.<sup>83</sup> Often the focus is on the social diversity of the panel, especially its racial<sup>84</sup> or gender<sup>85</sup> composition.<sup>86</sup>

Regardless of the specific social grouping under analysis, the studies tend to support Phillips’s claims about the value of informational diversity: Overall, they find that ‘minority’ judges on a panel can affect the choices of their colleagues in pertinent areas of the law because those judges possess information, experience, or expertise that is valuable to their colleagues. Research on US appellate judges, for example, demonstrates that the presence of a Black judge on an otherwise white panel, leads the white judges to issue decisions more favourable to Black plaintiffs in areas of the law where race is prominent (such as, voting rights and affirmative action).<sup>87</sup> Work on gender finds similar collegial effects. Panels with one of more female judges on international criminal tribunals give substantially longer sentences to sexually violent offenders than all-male panels (about 35 months longer),<sup>88</sup> and males serving on US appellate panels are significantly more likely to hold in favour of plaintiffs in gender-based employment cases when a female serves on the panel.<sup>89</sup> Also sitting comfortably with the importance of diverse perspectives are reports that gender- and racially-diverse panels more frequently consider alternative views on the questions presented in cases.<sup>90</sup>

### 3.2 Proposals for forward movement

Our review of the relevant studies is, admittedly, brief but for a good reason: Researchers have barely scratched the surface when it comes to using data to suss out diversity ‘bonuses’ on collegial courts. The basic question for our purposes is, what constitutes a ‘better’ judicial decision or a ‘better’ court? The list of possible benefits (forms of ‘better’) worthy of exploration is long but includes equity, legitimacy, quality, innovation, and creativity.

<sup>83</sup>J. Kastellec, ‘Panel Composition and Judicial Compliance on the U.S. Courts of Appeals’, (2007) 23 *Journal of Law, Economics and Organization* 421.

<sup>84</sup>E.g., Cox and Miles, *supra* note 21; Kastellec, *supra* note 55.

<sup>85</sup>E.g., Boyd, Epstein and Martin, *supra* note 14.

<sup>86</sup>Though of less relevance here, the panel’s ideological or partisan composition also has come under analysis. In general these studies, relative to research on race and gender, lean less on the effect of panel members on one another than on concerns about hierarchical superiors reversing the panel’s decision. For examples, see F. Cross and E. Tiller, ‘Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals’, (1998) 107 *Yale Law Journal* 2155; D. Beim and J. Kastellec, ‘The Interplay of Ideological Diversity, Dissents, and Discretionary Review in the Judicial Hierarchy: Evidence from Death Penalty Cases’, (2014) 76 *Journal of Politics* 1074; Kastellec, *supra* note 83.

<sup>87</sup>E.g., Cox and Miles, *supra* note 21; Kastellec, *supra* note 55; S. Farhang and G. Wawro, ‘Institutional Dynamics on the U.S. Court of Appeals: Minority Representation Under Panel Decision Making’, (2004) 20 *Journal of Law, Economics, and Organization* 299. For a review of these studies see Harris and Sen, *supra* note 13.

<sup>88</sup>K. King, J. Meernik and E. Kelly, ‘Deborah’s Voice: The Role of Women in Sexual Assault Cases at the International Criminal Tribunal for the Former Yugoslavia’, (2016) 98 *Social Science Quarterly* 548.

<sup>89</sup>Boyd, Epstein and Martin, *supra* note 14; Peresie, *supra* note 47.

<sup>90</sup>Haire, Moyer and Treier, *supra* note 30; R. Grey, K. McLoughlin and L. Chappell, ‘Gender and Judging at the International Criminal Court: Lessons from Feminist Judgment Projects’, (2021) 34 *Leiden Journal of International Law* 247.

The first sense of ‘better’ – *equity* – aligns with justifications for the rule of law: Are the decisions of collegial courts more impartial than those of individual judges? Studies of equity effects on collegial courts answer this question in the affirmative, suggesting that socially diverse courts may mitigate in-group bias. Recall, for example, the study showing that in family law disputes male judges tend to favour fathers and female judges favour mothers. The same study, though, reports that ‘these effects are dampened on panels that include judges of both genders’.<sup>91</sup> Likewise, a study of sentencing in the United States finds ‘that as the proportion of Black judges increases, white and Black judges are less likely to render incarceration sentences in cases with Black defendants and white judges are more likely to render incarceration sentences in cases with white defendants’.<sup>92</sup> Both studies, in other words, imply that social diversity can redress inequities caused by in-group bias. But far more work is needed, especially on courts outside the United States.

A second, though related, sense of ‘better’ is grounded in questions of the *legitimacy* of courts and how social diversity may produce possible legitimacy benefits. To be sure, there are many claims to the effect that ‘An all-male bench is no longer legitimate’<sup>93</sup> or that social-identity ‘representativeness is a democratic value that can serve to justify the [courts’] exercise of authority’.<sup>94</sup> Unfortunately, though, empirical support for these claims is limited – though not non-existent. In a notable experiment, researchers showed participants one of two articles: one reporting that Blacks comprise 23.2 per cent of the US federal bench and the other reporting that only 3.9 per cent of federal judges are Black.<sup>95</sup> Black respondents shown the ‘23.2 per cent’ article were more likely to respond positively to questions like ‘The courts can usually be trusted to make the right decision.’ Another experiment, this one on gender, found that the presence of women on a committee (perhaps akin to a judicial panel) conferred greater ‘institutional trust and acquiescence’ to its decisions.<sup>96</sup> The same experiment also demonstrated, intriguingly, that women’s presence had a legitimizing effect (especially on men) on decisions that ‘go against women’s interests’, such as decisions that question policies on sexual harassment.

Beyond these, empirical studies of the effect of social diversity on legitimacy are few and far between. Not even research specifically aimed at identifying the correlates of trust or confidence in judiciaries explicitly considers the diversity of the courts under analysis.<sup>97</sup>

With effort and creative thinking, this gap could be filled. Survey-experiments can be conducted (relatively cheaply and rapidly) using participants recruited via Amazon’s Mechanical Turk (MTurk). Although not without its share of problems, used with care MTurk can produce reasonably high-quality data.<sup>98</sup> Building the judges’ social identities into cross-national studies of

<sup>91</sup>Stribopoulos and Yahya, *supra* note 11.

<sup>92</sup>A. Harris, ‘Can Racial Diversity among Judges Affect Sentencing Outcomes?’, 2018, available at [www.allisonpharris.com/uploads/1/0/7/3/107342067/harris\\_diversitysentencing.pdf](http://www.allisonpharris.com/uploads/1/0/7/3/107342067/harris_diversitysentencing.pdf).

<sup>93</sup>S. Kenny, ‘Breaking the Silence: Gender Mainstreaming and the Composition of the European Court of Justice’, (2002) *Feminist Legal Studies* 257, at 257.

<sup>94</sup>Grossman, *supra* note 32. See also K. Hessler, ‘Women Judges or Feminist Judges?: Gender Representation and Feminist Values in International Courts’, *Journal of Social Philosophy* (forthcoming, available at [doi.org/10.1111/josp.12410](https://doi.org/10.1111/josp.12410)); A. Follesdal, ‘How Many Women Judges are Enough on International Courts?’, *Journal of Social Philosophy* (forthcoming, available at [doi.org/10.1111/josp.12399](https://doi.org/10.1111/josp.12399)); V. Hughes, ‘Adjudicating International Trade Cases in the World Trade Organization: Does Gender Make a Difference?’, in F. Baetens (ed.), *Identity and Diversity on the International Bench: Who is the Judge?* (2020).

<sup>95</sup>N. Scherer and B. Curry, ‘Does Descriptive Race Representation Enhance Institutional Legitimacy? The Case of the U.S. Courts’, (2010) 72 *Journal of Politics* 90. A control group was not given an article to read.

<sup>96</sup>A. Clayton, D. O’Brien and J. Piscopo, ‘All Male Panels? Representation and Democratic Legitimacy’, (2019) 63 *American Journal of Political Science* 113.

<sup>97</sup>E.g., P. Magalhães and N. Garoupa, ‘Judicial Performance and Trust in Legal Systems: Findings from a Decade of Surveys in over 20 European Countries’, (2020) 101 *Social Science Quarterly* 1743; M. Bühlmann and R. Kunz, ‘Confidence in the Judiciary: Comparing the Independence and Legitimacy of Judicial Systems’, (2011) 34 *West European Politics* 317.

<sup>98</sup>M. Chmielewski and S. Kucker, ‘An MTurk Crisis? Shifts in Data Quality and the Impact on Study Results’, (2020) 11 *Social Psychological and Personality Science* 464; R. Kennedy et al., ‘The Shape of and Solutions to the MTurk Quality Crisis’,

confidence, trust, and legitimacy is also possible, assuming substantive knowledge of the legal systems and of the relevant cleavages, along with careful data collection and analysis.

A third sense of 'better' refers to the ways in which collective decision making enhances the substantive quality of decisions. Questions of *quality*, as well as *innovation* and *creativity* directly implicate the benefits of diversity espoused by Phillips, Page, and other scholars<sup>99</sup> – and those benefits have been well documented both in the business world and in academia. Studies show that family-owned companies with at least 10 per cent female executives substantially outperform male-only companies by over 400 basis points per year;<sup>100</sup> that gender- and/or ethnically-diverse companies are significantly more likely to introduce innovations than firms dominated by one gender or ethnicity,<sup>101</sup> and that higher levels of racial diversity in firms are associated with more revenue, such that 'the mean revenues of organizations with low levels of racial diversity are roughly \$51.9 million, compared with . . . \$761.3 million for those with high levels of diversity'.<sup>102</sup> As for the academy: Freeman and Huang, among others, demonstrate that scientific papers authored by ethnically diverse teams are more consequential and impactful;<sup>103</sup> and that 'greater homophily is associated with publication in lower-impact journals and with fewer citations'.<sup>104</sup>

Although research running along similar lines has been conducted on juries,<sup>105</sup> studies of courts are almost non-existent.<sup>106</sup> We should rectify that. Just as scholars have used citation counts to measure the quality of papers produced by research teams that are socially diverse versus those that are homogeneous,<sup>107</sup> we could do the same with court decisions produced by socially diverse panels versus those that are not. Indeed, scholars have long used citations to judges' opinions to assess the judges' quality.<sup>108</sup> So too research might consider whether decisions issued by diverse teams of judges are less likely to be reversed by higher courts. Finally, projects on the relationship between social diversity and judicial innovation also could be adapted from existing research,

---

(2020) 8 *Political Science Research Methods* 614; E. Peer, J. Vosgerau and A. Acquisti, 'Reputation as a Sufficient Condition for Data Quality on Amazon Mechanical Turk', (2014) 46 *Behavior Research Methods* 1023.

<sup>99</sup>Phillips, *supra* note 81; Page, *supra* notes 71 and 78. For a review see A. Galinsky, 'Maximizing the Gains and Minimizing the Pains of Diversity: A Policy Perspective', (2015) 10 *Perspectives on Psychological Science* 742.

<sup>100</sup>Credit Suisse Research Institute, 'The CS Gender 3000 in 2019', available at [www.credit-suisse.com/about-us-news/en/articles/news-and-expertise/csgender-3000-report-2019-201910.html](http://www.credit-suisse.com/about-us-news/en/articles/news-and-expertise/csgender-3000-report-2019-201910.html).

<sup>101</sup>C. Østergaard, B. Timmermans and K. Kristinsson, 'Does a Different View Create Something New? The Effect of Employee Diversity on Innovation', (2011) 40 *Research Policy* 500; P. Parrotta, D. Pozzoli and M. Pytlíkova, 'The Nexus between Labor Diversity and Firm's Innovation', (2014) 27 *Journal of Population Economics* 303. See also C. Dezső and D. Ross, 'Does Female Representation in Top Management Improve Firm Performance? A Panel Data Investigation', (2012) 33 *Strategic Management Journal* 1072.

<sup>102</sup>A similar pattern holds for gender-diverse firms. C. Herring, 'Does Diversity Pay?: Race, Gender, and the Business Case for Diversity', (2009) 74 *American Sociological Review* 208.

<sup>103</sup>R. Freeman and W. Huang, 'Collaborating with People Like Me: Ethnic Coauthorship within the United States', (2015) 33 *Journal of Labor Economics* S289. See also B. AlShebli, T. Rahwan and W. Woon, 'The Preeminence of Ethnic Diversity in Scientific Collaboration', (2018) 9 *Nature Communications* 5163; M. Nielsen et al., 'Opinion: Gender Diversity Leads to Better Science', (2017) 114 *Proceedings of the National Academy of Sciences* 1740.

<sup>104</sup>Freeman and Huang, *ibid.*, at S289.

<sup>105</sup>See, e.g., S. Sommers, 'On Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations', (2006) 90 *Journal of Personality and Social Psychology* 597 (finding that diverse mock juries exchanged 'a wider range of information than all-white groups').

<sup>106</sup>Exceptions include a study showing that gender-diverse US appellate panels were more likely to innovate in the area of sexual harassment. L. Moyer and H. Tankersley, 'Judicial Innovation and Sexual Harassment Doctrine in the US Courts of Appeals', (2012) 65 *Political Research Quarterly* 784. See also Haire, Moyer and Treier, *supra* note 30.

<sup>107</sup>Freeman and Huang, *supra* note 103. See also G. Caldeira, 'On the Reputation of State Supreme Courts', (1983) 5 *Political Behavior* 83.

<sup>108</sup>M. Gulati et al., 'Judging Women', (2011) 8 *Journal of Empirical Legal Studies* 504; R. Anderson, 'Distinguishing Judges: An Empirical Ranking of Judicial Quality in the United States Court of Appeals', (2011) 76 *Missouri Law Review* 315; R. Smyth and M. Bhattacharya, 'What Determines Judicial Prestige? An Empirical Analysis for Judges of the Federal Court of Australia', (2003) 5 *American Law and Economics Review* 233; W. Landes, L. Lessig and M. Solimine, 'Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges', (1998) 27 *Journal of Legal Studies* 271.

including work that traces the development of a ‘new’ rule or standard and tracks its diffusion (or lack thereof) across courts and societies.<sup>109</sup>

#### 4. From promise to reality: Diversity is the key

Our primary goal in this essay was to offer what we think (hope!) are promising avenues for future work on social identity and social diversity. But, as we have signaled throughout, the studies will not write themselves; hard work and creativity are required to move in the proposed directions. That work is likely to come in many different forms, using data developed from in-depth interviews, structured surveys, experiments, court records, and so on. Moreover, researchers may well confront obstacles along the way, such as preserving the privacy of litigants and judges.<sup>110</sup>

We are agnostic about the form of the research; data are data and methods are methods.<sup>111</sup> Which types get selected (should) depend on the researchers’ questions and goals.

Where we feel far more strongly is in how scholars go about building collaborations. After all, an important implication of the research we have reviewed is that scholars should develop diverse research teams. Relevant substantive diversity, of course, is crucial. A survey researcher, a specialist in judicial behaviour, and an expert in law are more likely to generate higher quality work products than, say, a team of all survey researchers or of all specialists in judicial behaviour. But, as Phillips suggests, a male Norwegian survey researcher and a female Taiwanese survey researcher are likely to bring different perspectives to the table with similarly beneficial results.

Building social diverse teams may not be easy, but the lesson from decades’ worth of research is that the benefits may well outweigh the costs. To once again quote Phillips, ‘The pain associated with diversity can be thought of as the pain of exercise. The pain . . . produces the gain. In just the same way . . . we need diversity if we are to change, grow and innovate.’<sup>112</sup> This holds for all teams and all organizations – courts not excepted.

**Acknowledgements.** Prepared for the iCourts Symposium, ‘How Personal Characteristics of International Judges Affect Their Rulings: Socio-Legal Perspectives’, in the *Leiden Journal of International Law*.

We thank the editors and an anonymous reviewer for their very helpful suggestions. Epstein also thanks the John Simon Guggenheim Foundation, the National Science Foundation, and Washington University for supporting her research on judicial behaviour.

<sup>109</sup>Moyer and Tankersley, *supra* note 106; H. Glick, ‘Judicial Innovation and Policy Re-Invention: State Supreme Courts and the Right to Die’, (1992) 45 *Western Political Quarterly* 71; G. Caldeira, ‘The Transmission of Legal Precedent: A Study of State Supreme Courts’, (1985) 79 *American Political Science Review* 178; R. Bird and D. Smythe, ‘Social Network Analysis and the Diffusion of the Strict Liability Rule for Manufacturing Defects, 1963–87’, (2012) 37 *Law & Social Inquiry* 565.

<sup>110</sup>An extreme example was imposed in France in 2019 (Art. 33 of the Justice Reform Act). The French legislation bans the publication, evaluation, analysis, comparisons or predictions of the behaviour of individual judges. Were researchers to release datasets revealing the judges’ names, they could face a maximum sentence of five years in prison.

<sup>111</sup>J. Patty, ‘Responding to a Petition to Nobody (or Everybody)’, *Math of Politics*, 6 November 2015, available at [www.mathofpolitics.com/2015/11/06/responding-to-a-petition-to-nobody-or-everybody/](http://www.mathofpolitics.com/2015/11/06/responding-to-a-petition-to-nobody-or-everybody/).

<sup>112</sup>Phillips, *supra* note 81.