
INTELLECTUAL HUMILITY BY DESIGN: HOW COURTS EARN LEGITIMACY AND IMPROVE DECISIONS

Steve Leben *

I. INTRODUCTION

Courts decide disputes as one branch of our government. Their authority depends not only on getting the outcome right but also on public acceptance of their authority. In a time when confidence in institutions is under pressure, courts cannot assume that legitimacy will take care of itself. They must earn it—case by case—through both the substance of what they decide and the way they exercise authority in public view.

A large body of procedural-justice research explains one side of that equation. People are more willing to accept court decisions, even unfavorable ones, when the process communicates voice, neutrality, respect, and trustworthy motives.¹ At the same time, courts are not legitimacy machines alone. They decide matters of profound consequence, and the public expects them to get it right. Decision science helps explain

* Steve Leben is the Douglas R. Stripp Distinguished Professor of Law at the University of Missouri–Kansas City School of Law. He was a judge on the Kansas Court of Appeals from 2007 to 2020 and a Kansas trial judge from 1993 to 2007. Author’s Note: I greatly appreciate the collaborations throughout my career with Kevin Burke, with whom I have written and lectured about many of the topics covered here for almost two decades; Tom Tyler, whose work has made most of mine possible; Emily LaGratta, with whom I’m now collaborating along with Kevin Burke on the proceduraljustice.org website; David Rottman, who helped us in adapting this work for court use and with the original procedural-justice website; and Nancy Levit, who has been the mentor I needed as I moved from the judiciary to the academy in 2020 and beyond. I also appreciate the help of my student research assistant, Sophie Green. Although all of them have impacted my work, none of them are responsible for any errors that may be found in this Article.

1. Tom R. Tyler, *Procedural Justice, Legitimacy, and the Effective Rule of Law*, 30 CRIME & JUST. 283, 284, 293-94, 298, 300, 350 (2003) [hereinafter *Procedural Justice, Legitimacy, and the Effective Rule of Law*]; Tom R. Tyler, *Whither Legitimacy? Legal Authority in the Twenty-First Century*, 19 ANN. REV. L. SOC. SCI. 14.1, 14.5-.6 (2023) [hereinafter *Whither Legitimacy*].

why that aspiration is difficult even for conscientious judges.² Like all human judgment, judicial decision-making is vulnerable to predictable errors—overconfidence, framing effects, motivated reasoning, and unwanted variability (“noise”) that can make outcomes feel like luck of the draw.³

This Article argues that a third research tradition—intellectual humility—can connect those two concerns. Intellectual humility is the recognition that one’s knowledge may be incomplete and one’s conclusions fallible, paired with a disciplined openness to reasons and evidence that cut against an initial view.⁴ Research in psychology links intellectual humility to behaviors that matter for justice-system leaders: greater willingness to listen, greater openness to opposing views, and reduced “myside” distortions, among others.⁵ Importantly for courts, intellectual humility is not merely a moral aspiration. Social psychologists have developed measures for it and have begun identifying practices that can cultivate it.⁶

The Article’s contribution is in combining and applying these areas together. Building on legal scholarship that emphasizes humility—especially intellectual humility—as an input-side discipline for judging,⁷ this Article weaves procedural

2. See Jeffrey J. Rachlinski & Andrew J. Wistrich, *Gains, Losses, and Judges: Framing and the Judiciary*, 94 NOTRE DAME L. REV. 521, 569-70, 573 (2019); Jeffrey J. Rachlinski, Andrew J. Wistrich, & Chris Guthrie, *Can Judges Make Reliable Numeric Judgements? Distorted Damages and Skewed Sentences*, 90 IND. L. REV. 695, 698 (2015); Jeffrey J. Rachlinski, Andrew J. Wistrich, & Chris Guthrie, *Altering Attention in Adjudication*, 60 UCLA L. REV. 1586, 1615-16 (2013) [hereinafter *Altering Attention*].

3. Andrzej Uhl & Justin T. Pickett, *Noise in Judicial Decision-Making: A Research Note*, 64 CRIMINOLOGY 195, 196, 203 (2026); see Andrew J. Wistrich, Chris Guthrie, & Jeffrey J. Rachlinski, *Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding*, 153 U. PA. L. REV. 1251, 1323-24 (2005); see Rachlinski & Wistrich, *supra* note 2, at 570.

4. MARK R. LEARY, *THE PSYCHOLOGY OF INTELLECTUAL HUMILITY* 3-4 (2018), [https://perma.cc/9ZU4-MYQU].

5. *Id.* at 8-11; Elizabeth J. Krumrei-Mancuso et al., *Toward an Understanding of Collective Intellectual Humility*, 29 TRENDS COGNITIVE SCIS. 15, 15 (2025); Shauna M. Bowes et al., *Stepping Outside the Echo Chamber: Is Intellectual Humility Associated with Less Political Myside Bias?*, 48 PERSONALITY & SOC. PSYCH. BULL. 150 (2022).

6. Krumrei-Mancuso et al., *supra* note 5, at 19, 21.

7. See CHAD M. OLDFATHER, *JUDGES, JUDGING, AND JUDGMENT: CHARACTER, WISDOM, AND HUMILITY IN A POLARIZED WORLD* 186-89, 201-05 (2024) (distinguishing humility as an “output-side” restraint from “input-side” intellectual humility and urging emphasis on judicial character and norms amid legitimacy pressures); Katie Miller, *Intellectual Humility: A Necessary Precondition to Building Trust in Courts*, 12 INT’L J. CT.

justice, decision science, and intellectual humility into a single framework tailored to the everyday work of courts—and it translates that framework into concrete routines for trial courts, appellate courts, chambers, and court governance.

The core claim here is that weaving these three strands together produces a framework stronger than any one standing alone. Procedural justice identifies what makes authority acceptable. Decision science identifies why error and inconsistency remain predictable risks. Intellectual humility helps bridge the two by making error-checking and fair engagement feel like professional method rather than personal weakness. Put differently, intellectual humility can help courts adopt and sustain both procedural-fairness practices and decision-hygiene safeguards—especially in the settings where judging actually happens: courtroom interaction, chambers deliberation, decision announcements, and court governance.

The Article proceeds in three steps. Parts II–IV summarize the procedural-justice, decision-science, and intellectual-humility literatures in terms relevant to judges and courts.⁸ Part V integrates them into an “Intellectual Humility by Design” model and operationalizes that model for trial courts, appellate courts, chambers, and court administration, with reasons and explanations treated as the bridge between decision quality and legitimacy.⁹ Part VI provides practical tools and adoption pathways, along with a measurement agenda and candid limits.¹⁰ The aim is to offer courts an evidence-informed way to earn legitimacy while improving the reliability of their judgments, through routines that are realistic, teachable, and durable.¹¹

ADMIN. 4, 4-5 (arguing that intellectual humility is a necessary condition for building trust in courts and “seek[ing] to build a bridge between” external accountability and the intentions of court leaders).

8. *See infra* Parts II–IV.

9. *See infra* Part V.

10. *See infra* Part VI.

11. *See infra* Section VI.A.

II. PROCEDURAL JUSTICE AND JUDICIAL LEGITIMACY

Courts depend on public acceptance of decisions, and perceived legitimacy is the mechanism that makes lawful authority durable in a democracy.¹² A central claim of procedural-justice literature is that compliance with legal authorities is best obtained when people feel that authorities are legitimate, not simply because people fear sanction.¹³ Fostering a sense that courts are legitimate seems especially important in a time of a divided political climate and declining trust generally in institutions.¹⁴

From a procedural-justice viewpoint, legitimacy is more than agreement with the end result: it is shown in a felt obligation to defer to and comply with decision-makers combined with a belief that legal authorities share (or at least respect) community values.¹⁵ While getting results right and finding the truth matter,¹⁶ popular legitimacy is most strongly shaped by how courts exercise their authority.¹⁷ So Part II uses procedural justice as this

12. Tom R. Tyler & Justin Sevier, *How Do The Courts Create Popular Legitimacy?: The Role of Establishing the Truth, Punishing Justly, and/or Acting Through Just Procedures*, 77 ALB. L. REV. 1095, 1096-97 (2013-2014).

13. See generally TOM R. TYLER, WHY PEOPLE OBEY THE LAW 165, 167, 170, 172, 175, 177-78 (2006); *Procedural Justice, Legitimacy, and the Effective Rule of Law*, *supra* note 1, at 284, 290, 297, 320, 322-23.

14. See Steve Leben, *Partisan Judges*, 77 ARK. L. REV. 517, 536-39 (2024) [hereinafter *Partisan Judges*].

15. See Thomas C. O'Brien, Tom R. Tyler, & Tracey L. Meares, *Building Popular Legitimacy with Reconciliatory Gestures and Participation: A Community-Level Model of Authority*, 14 REGUL. & GOVERNANCE 821, 824-25, 834-35 (2020); Tom R. Tyler & Jonathan Jackson, *Popular Legitimacy and the Exercise of Legal Authority: Motivating Compliance, Cooperation, and Engagement*, 20 PSYCH., PUB. POL'Y, & L. 78, 79-81, 86, 90 (2014); *Whither Legitimacy*, *supra* note 1, at 14.5; see generally TYLER, *supra* note 13, at 170, 172.

16. In addition to the components of procedural justice, which are discussed below in text, Tyler and Sevier found that the public's view of court legitimacy was also shaped by a sense that courts sentence justly and establish truth. See Tyler & Sevier, *supra* note 12, at 1098.

17. See *id.* at 1125 (noting that in a survey of the general public, "the primary judgment shaping popular legitimacy is the procedural justice of court actions"); *Whither Legitimacy*, *supra* note 1, at 14.3-.4, 14.6 (explaining that legitimacy as a basis for deference to legal authorities and procedural justice as a key antecedent to legitimacy have been established in multiple studies); Tom R. Tyler, *Procedural Justice and the Courts*, 44 CT. REV. 26, 28, 31 (2007) [hereinafter *Procedural Justice and the Courts*] (noting that in a careful study of

Article's first building block: a legitimacy-centered explanation of why process matters.¹⁸ It will also serve as a foundation for later arguments about how intellectual humility—treated as a series of practices rather than a moral trait—can strengthen both acceptance of decisions (discussed here) and decision reliability (discussed in Part III).¹⁹

Procedural-justice researchers have looked at several potential reasons people might accept court decisions. They might be most influenced by the substantive result—whether the court got it right.²⁰ They might be most influenced by whether the court ruled in their favor.²¹ Or they might be most influenced by whether the court handled the case through fair procedures.²² Over several decades of research, fair procedures have been established as the most important of those three potential influences.²³

So, what does the public consider fair procedures to be? Yale law professor Tom Tyler and those who have worked with him have consistently identified four components: voice (a meaningful opportunity to be heard), neutrality (rule-based, unbiased decision-making that is transparent enough for that to be seen), respect (dignified treatment and explicit regard for rights), and trustworthiness (signs that the decision-maker is sincere, caring, and trying to do what is right rather than acting out of self-interest or prejudice).²⁴ These elements are analytically distinct (and generally have statistically significant effects

California court participants, “the primary factor shaping the willingness to accept decisions was the fairness of court procedures” and that “[t]he willingness to accept court decisions . . . was about the procedures used to reach those decisions, not the decisions themselves.”).

18. See *supra* notes 12-41 and accompanying text.

19. See *infra* Part III.

20. See Tyler & Sevier, *supra* note 12, at 1096.

21. See TYLER, *supra* note 12, at 5-6; See *Procedural Justice, Legitimacy, and the Effective Rule of Law*, *supra* note 1, at 288-89, 294, 313.

22. See Tyler & Sevier, *supra* note 12, at 1096.

23. See *Partisan Judges*, *supra* note 14, at 555; Steve Leben, *An Expectation of Empathy*, 51 WASHBURN L.J. 49, 51-52 (2011) [hereinafter *Expectation of Empathy*]; Kevin Burke & Steve Leben, *Procedural Fairness: A Key Ingredient in Public Satisfaction*, 44 CT. REV. 4, 5 (2008) [hereinafter *Public Satisfaction*]; Kevin Burke & Steve Leben, *Procedural Fairness*, in ENHANCING JUSTICE: REDUCING BIAS 229, 230, 232 (2017).

24. See Tyler & Sevier, *supra* note 12, at 1105-07; see *Procedural Justice and the Courts*, *supra* note 17, at 30-31.

individually),²⁵ but they are not independent in practice. Courts earn legitimacy by producing an integrated experience of being heard and treated fairly by an authority whose reasoning appears principled and sincere.²⁶

These elements are critical to judicial performance,²⁷ so let's consider each one in a bit more detail. Voice is more than just letting people talk. People do want an opportunity to tell their story in their own words, but they also want to believe that an authority figure has genuinely listened and considered what they said before deciding.²⁸ Tyler emphasizes that voice matters in part because it shapes the other procedural-justice judgments people make—neutrality, trust, and respect—by signaling that the authority is not merely processing cases but is genuinely engaging with people.²⁹ That connection is especially important in high-volume court proceedings, where parties may infer (rightly or wrongly) that speed and routine processing mean that they were not really heard.

Neutrality is the perception that decisions are made by consistently applied principles rather than personal opinions or bias.³⁰ To gain perceived legitimacy, neutrality must be visible to those subject to the decision. Tyler and Sevier stress that neutrality therefore “requires an explanation for the process and accountability in terms of the rules used to make decisions.”³¹ Explanations that connect outcomes to rules help people experience the court as principled, not arbitrary.³² In this sense, neutrality is closely connected to explanation: rule-based

25. See *id.* at 30-31 (finding direct effects for neutrality, respect, and trust and indirect effects, through the other factors, for voice); see Tyler & Sevier, *supra* note 12, at 1126 (finding all four elements—voice, neutrality, respect, and trust—“distinctly important” in public acceptance of court legitimacy).

26. See Tyler & Sevier, *supra* note 12, at 1106-07.

27. David B. Rottman & Tom R. Tyler, *Thinking About Judges and Judicial Performance: Perspective of the Public and Court Users*, 4 OÑATI SOCIO-LEGAL SERIES 1046, 1049-50 (2014).

28. See *Procedural Justice and the Courts*, *supra* note 17, at 30.

29. *Id.* at 31.

30. *Id.* at 30.

31. See Tyler & Sevier, *supra* note 12, at 1105-06.

32. See *Procedural Justice and the Courts*, *supra* note 17, at 30-31.

decision-making becomes visible through reasons that show what standards were applied and how they controlled the result.³³

Respect is more than politeness in open court. The judge is a status figure representing the government.³⁴ Because of that, Tyler and Sevier explain, the quality of treatment provided by the judge is “a statement about status” in society of the people appearing before the court.³⁵ People are sensitive to whether they are treated with dignity and with respect for their rights. Even when a decision is legally correct, dismissive or demeaning treatment can undermine legitimacy because it communicates that the person is not given full status in society. Explaining a person’s rights, even when ruling against them, also shows respect. This respectful treatment affirms that people are valued and that their concerns and rights have been taken seriously, not swept aside without good reason.

The fourth element—trustworthiness—captures an important feature people look for in a decision-maker with authority over them. People draw inferences about the character of legal authorities: whether they are sincere, honest, caring, and acting in good faith rather than out of bias or self-interest.³⁶ Tyler emphasizes that these judgments are often based on whether people believe court personnel are listening to and considering their views, are being open about the basis for their actions, and are trying to do what is right for everyone involved.³⁷ As Tyler and Sevier note, authorities communicate trustworthy motives when they listen and “explain or justify their actions in ways that show an awareness of and sensitivity to people’s needs and concerns.”³⁸ This is legitimacy in the everyday sense—when wielding power, the court exercises that power conscientiously, not simply because it can.

These four elements are interrelated, and judges often advance several at once through a single practice: good

33. See Tyler & Sevier, *supra* note 12, at 1105-06.

34. *Id.* at 1106.

35. See Tyler & Sevier, *supra* note 12, at 1106; see also *Procedural Justice and the Courts*, *supra* note 17, at 30-31.

36. See Tyler & Sevier, *supra* note 12, 1106-07.

37. See *Procedural Justice and the Courts*, *supra* note 17, at 31.

38. See Tyler & Sevier, *supra* note 12, at 1106-07.

explanations.³⁹ A well-crafted explanation (1) demonstrates voice by making it clear that the decision-maker heard and considered the parties' key points; (2) shows neutrality by revealing a rule-based path from governing standards and facts to the result; and (3) supports trustworthiness by communicating good faith—that the judge is not just acting on whim or bias but is sincerely trying to do what is right within the constraints of applicable legal standards.⁴⁰ Explanations, then, are not only an issue of writing craft. They are also one of the best ways to operationalize procedural justice in day-to-day court work—a point that becomes central in Part V's Intellectual Humility by Design framework.⁴¹

In sum, procedural-justice research explains how courts can earn legitimacy through fair process: when people experience voice, neutrality, respectful treatment, and trustworthy motives, they are more likely to accept decisions—even adverse ones. Since courts decide matters of consequence—like termination of parental rights, the scope of constitutional protections, and the length of incarceration—legitimacy is understandably weakened when outcomes appear arbitrary or error-prone. The question, then, is how courts can pursue both aims at once: fair process and reliable judgment under the ordinary constraints of adjudication—limited information, time pressure, heavy caseloads, emotionally charged disputes, and repeated exposure to contested narratives. The next Part turns to decision science, which identifies predictable vulnerabilities in human judgment and the kinds of structural safeguards that can reduce their impact.

39. *Id.* at 1106-07, 1125 n.80 (noting the specific value of explanations for neutrality and trustworthiness, as well as the interrelatedness of the four elements); see *Procedural Justice and the Courts*, *supra* note 17, at 30-31 (noting how explanations support views of neutrality).

40. See Tyler & Sevier, *supra* note 12, at 1105-07 (explaining that trustworthiness cues are communicated when authorities listen and explain and that neutrality requires explanations and accountability); see also *Procedural Justice and the Courts*, *supra* note 17, at 31 (noting that voice indirectly shapes neutrality, trust and respect, and that trustworthiness involves perceptions of listening and openness).

41. See *infra* Section V.C.

III. DECISION SCIENCE AND PREDICTABLE JUDICIAL ERROR

There is a famous quotation in the statutory-interpretation field from Professors Henry Hart and Albert Sachs about how we should think about legislators when figuring out what a statute was intended to do. “[A]ssume,” they said, “unless the contrary unmistakably appears, that the legislature was made up of reasonable persons pursuing reasonable purposes reasonably.”⁴² In the case of legislators, we may sometimes think that legislators have other motives. Even so, it often seems advisable to make that assumption when attributing a purpose to a statute.

In the judicial branch, judges generally are trying to get the answer right.⁴³ But even well-intentioned conscientious judges are subject to the same mental errors the rest of us are. And like most of us, judges face pressures that can easily affect thought processes—stress, uncertainty, time pressure, and heavy dockets. For judges and the rest of us, judgment error often has two components: bias and noise.⁴⁴

Bias is a systematic directional distortion.⁴⁵ It need not be the sort of pejorative bias that we think of with respect to racial bias. If, for example, hearing inadmissible information leads most judges to a different ruling on the merits—even though they theoretically have factored out the inadmissible evidence—that is

42. HENRY M. HART, JR. & ALBERT M. SACKS, *THE LEGAL PROCESS: BASIC PROBLEMS IN THE MAKING AND APPLICATION OF LAW* 1378 (William N. Eskridge Jr. & Philip P. Frickey, eds. 1994).

43. Birte Englich, Thomas Mussweiler, & Fritz Strack, *Playing Dice with Criminal Sentences: The Influence of Irrelevant Anchors on Experts' Judicial Decision Making*, 32 PERSONALITY & SOC. PSYCH. BULL. 188, 198 (2006) (noting that judges are professionally motivated to be accurate). The author's experience in twenty-seven years as a judge also supports the view that most judges are trying to get the right answer most of the time.

44. DANIEL KAHNEMAN, OLIVIER SIBONY, & CASS R. SUNSTEIN, *NOISE: A FLAW IN HUMAN JUDGMENT* 4-7 (2021) (distinguishing “[b]ias and noise—systematic deviation and random scatter” as different components of judgment error); see Uhl & Pickett, *supra* note 3, at 196 (defining noise in sentencing as unwanted and unsystematic variation leading to different sentences for similarly situated defendants).

45. Jeff Aronson, *A Word About Evidence: 6. Bias – A Proposed Definition*, CATALOGUE OF BIAS, [<https://perma.cc/PMC2-8RF3>] (last visited Mar. 13, 2026).

a systematic distortion of the result in a specific direction.⁴⁶ That's bias.

A separate problem is unwanted variability that may not be in a single direction—what Daniel Kahneman, Olivier Sibony, and Cass Sunstein call noise.⁴⁷ At least in theory, for example, a study of 100 sentencing decisions might find no perceptible bias based on averages. But if there is great variability, some defendants would rightly wonder why their sentence had been so much higher than another who was similarly situated. Too much variability—noisy decision-making—can be a threat to rule-of-law values. No one wants their experience with the judicial system to be a “luck-of-the-draw” experience.⁴⁸

But noise is hard to eliminate: “wherever there is judgment, there is noise[.]”⁴⁹ For that reason, many organizations study it through noise audits that compare judgments on matched or standardized case samples.⁵⁰ Consider an insurance company that asked multiple underwriters and claims adjusters to evaluate the same written case files.⁵¹ Executives expected modest variation, but the audit revealed far larger variability, essentially turning case assignment into the luck-of-the draw experience no one would rationally design a system to provide.⁵²

46. See Wistrich, Guthrie, & Rachlinski, *supra* note 3, at 1292, 1297, 1303, 1307, 2005 (reporting experiments finding that relevant but inadmissible information influenced judges' decisions); *Altering Attention*, *supra* note 2, at 1596-97 (summarizing studies showing that shifting what is salient to judges can alter outcomes, including judges' willingness to ignore inadmissible evidence).

47. See Uhl & Pickett, *supra* note 3, at 196 (discussing noise in sentencing); KAHNEMAN, SIBONY & SUNSTEIN, *supra* note 43, at 43-54 (explaining how noise can affect matters of judgment).

48. See Uhl & Pickett, *supra* note 3, at 197, 202 (documenting noise in simulated sentencing decisions by early-career judges in Poland, concluding that “punishment was more like roulette than just deserts”); KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 43, at 69-78 (summarizing a noise audit of U.S. federal criminal sentences finding considerable unexplained variance and concluding that “we suspect that the amount of noise defendants face in actual courtrooms is even larger than what we see here”).

49. KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 12.

50. *Id.* at 25-26 (defining a “noise audit” as having many people evaluate the same case and making variability visible).

51. *Id.*

52. *Id.* at 26-27 (reporting the insurance company noise-audit results; median underwriting difference fifty-five percent, claims-adjuster median difference forty-three percent).

In response to this excess variability, Kahneman, Sibony, and Sunstein suggest several measures that can quiet the noise.⁵³ They begin with a practical premise: organizations often cannot improve what they do not measure.⁵⁴ A noise audit makes inconsistency visible and creates institutional pressure to respond—through training, revised decision frameworks, and structured review—while also allowing the organization to test whether the reforms helped to reduce the noise.⁵⁵ The point is to narrow the range of outcomes for similar cases while preserving discretion where it makes sense to apply it.⁵⁶ That translates well to courts, where discretion is inevitable. Ideally, though, unexplained variability could be reduced through clearer factor frameworks, more consistent explanation practices, and routine feedback.

Criminal sentences are one of the areas where judicial decisions have been subjected to noise audits. In an early large-scale exercise, federal judges were asked to impose sentences in hypothetical cases summarized in identical presentence reports.⁵⁷ The resulting dispersion was striking—both in the length of recommended incarceration and, in many cases, in whether incarceration was appropriate at all.⁵⁸ The sentencing-guidelines movement that followed can be understood, in part, as an institutional attempt to reduce that excess variability by constraining discretion through structured ranges tied to offense seriousness and criminal history.⁵⁹ Empirical work during the guidelines era reported meaningful reductions in inter-judge disparity, even as critics argued that the resulting system risked

53. See generally *id.* at 9, 25657, 275 (introducing “decision hygiene” and summarizing noise-reduction strategies including sequencing information, aggregating individual judgments, and guidelines).

54. KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 53-54.

55. *Id.* at 53-54, 221.

56. *Id.* at 43-44, 374-76 (discussing considerations about how much noise should be acceptable given other factors).

57. *Id.* at 15-16.

58. *Id.* at 14-21 (reporting federal sentencing studies, including one large-scale study in 1974 in which fifty federal judges in identically hypothetical cases to sentences that had an “astounding” dispersion).

59. KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 17-18.

unacceptable mechanization and insufficient attention to individual circumstances.⁶⁰

A recent study in Poland underscores that the underlying problem is not confined to a particular period or legal culture: Andrzej Uhl and Justin Pickett describe a naturally occurring noise audit in Poland in which early career judicial candidates, working from a uniform court file exceeding one hundred pages in a high-stakes examination, produced widely dispersed sentencing outcomes.⁶¹ Identical defendants could receive punishments ranging from modest financial penalties to substantial prison terms.⁶² Sentencing noise is real and measurable, and institutional choices—including guidelines and other forms of decision structuring—can reduce unjustified dispersion. While this Article is not focused specifically on correcting sentencing disparities, later sections will review potential ways of reducing noise, whether in sentencing or other judicial decisions.

As already noted, noise is not the only cognitive vulnerability found among judges. Before turning to empirical studies about other cognitive errors among judges, though, a brief point about how people think and make decisions is helpful. We are bombarded by sensory information, and we can't take it all in. So much of what we do to process information happens behind the scenes, without conscious thought.⁶³ Most of our information processing is reflexive and automatic—what Kahneman popularized as “fast” thinking.⁶⁴ Much of the time it works well. It relies on schemas and other learned mental patterns that allow recognition and rapid response without deliberation from scratch.⁶⁵ You see a stop sign and stop; you don't need to think

60. *Id.* at 18-19 (finding significant success of guidelines in reducing disparities but combined with fairness criticisms).

61. Uhl & Pickett, *supra* note 3, at 197, 199.

62. *Id.* at 199-204 (finding “a large amount of noise” in judicial sentencing decisions in a carefully designed experiment with candidates completing judicial training in Poland).

63. Pamela Casey, Kevin Burke, & Steve Leben, *Minding the Court: Enhancing the Decision-Making Process*, 49 CT. REV. 76, 76-77 (2013) (explaining reflexive versus reflective processing and the use of schemas and heuristics under limited reflective capacity).

64. DANIEL KAHNEMAN, THINKING, FAST AND SLOW 19-49 (2011) (explaining mental processing and introducing Kahneman's classification of fast and slow thinking).

65. *Id.*

about it. A firefighter at the site of a fire quickly goes to work without stopping to analyze every possible option; training and experience supply a rapid first impression of what to do next.⁶⁶

But fast processing has predictable failure modes. Schemas and shortcuts that promote speed aren't always accurate, and they can be overgeneralized or misapplied to atypical cases.⁶⁷ People may fuse events that often occur together into a causal story even when the relationship is coincidental. In that sense, implicit bias can often be understood as an automatically activated association that may reflect incomplete or distorted learning rather than careful case-specific evaluation.⁶⁸

When the initial impression is unreliable, the decision-maker may benefit from what Kahneman calls “slow” thinking—taking more time to test first impressions against a governing standard, to search deliberately for disconfirming evidence, and to consider alternate explanations.⁶⁹ That is the backdrop for research on judicial decision-making.⁷⁰ Mostly in controlled settings, researchers have sought to determine whether judges—despite their professional commitment to deciding accurately and fairly—show predictable susceptibility to common cognitive

66. Gary Klein, Roberta Calderwood, & Anne Clinton-Cirocco, *Rapid Decision Making on the Fire Ground: The Original Study Plus a Postscript*, 4 J. COGNITIVE ENG'G & DECISION MAKING 186, 194 (2010) (finding in a study of firefighters that in most cases, “[o]ptions were selected without any reports of conscious examination, evaluation, or analysis [T]he events triggered an immediate cognition of what had to be done, and action was taken”); see KAHNEMAN, *supra* note 64, at 236-40 (summarizing and explaining Klein's firefighter research); Casey, Burke, & Leben, *supra* note 63, at 80 (same).

67. *Heuristics and Schemas: Mental Shortcuts and Frameworks*, PROFESSOR RJ STARR, [https://perma.cc/D7V7-A623] (last visited Mar. 1, 2026).

68. See JENNIFER K. ELEK & ANDREA L. MILLER, NAT'L CTR. FOR STATE CTS., *THE EVOLVING SCIENCE ON IMPLICIT BIAS*, 4-5 (2021) (explaining that implicit bias involves automatic associations shaped by social learning and environment rather than deliberative evaluation); Casey, Burke, & Leben, *supra* note 63, at 80-81 (explaining that schemas can mislead when people infer causation from co-occurrence “as in superstitions”).

69. See KAHNEMAN, *supra* note 64, at 417 (urging decision-makers to “recognize the signs that you are in a cognitive minefield, slow down, and ask for reinforcement from System 2[.]” which he calls slow thinking); Casey, Burke, & Leben, *supra* note 63, at 79, 87 (describing dual-process models and contrasting a “reflexive” system with a “reflective,” “slow,” conscious system).

70. See generally Andrew J. Wistrich, *Secret Shoals of the Shadow Docket*, 23 NEV. L.J. 863, 875-77, 884-86 (2023) (reviewing past studies of judicial cognition errors); Casey, Burke, & Leben, *supra* note 63, at 79-83 (summarizing dual-process decision-making, including cognitive heuristics and implicit bias as predictable error sources for judges).

errors and under what conditions reflective reasoning might reduce such errors.⁷¹

Unsurprisingly, judges show the same cognitive-error biases others do. Anchoring is a well-studied example.⁷² For judges, anchoring is the influence a number can have on a legal judgment even when the number has no legitimate relationship to the legal task.⁷³ Several studies have shown that once a number is introduced into the discussion and thus the judge's mind—whether from a demand, a guideline, a charging decision, an initial recommendation, or a legally irrelevant prompt—it can shape the eventual sentence or decision by pulling the judge's estimate toward the anchor.⁷⁴

A study of German judges shows how significant an irrelevant number can be. Thirty-seven judges got a set of materials for the sentencing of a twelve-time shoplifter.⁷⁵ Before giving their sentencing recommendation, they were also given the prosecutor's recommendation—but they were explicitly told that the prosecutor's recommendation was “randomly determined” and “therefore, it does not reflect any judicial expertise[.]”⁷⁶ Yet with half the group given a prosecutor's recommended probation of nine months and half of three months, the sentences for the high-anchor group were fifty percent higher than for the low-anchor group.⁷⁷ In another study by the same researchers, the anchoring effect was seen from the prosecutor's recommended

71. See Wistrich, *supra* note 70, at 875, 884-86 (surveying judicial-cognition-error research and its implications for courts).

72. See Chris Guthrie, Jeffery J. Rachlinski, & Andrew J. Wistrich, *Inside the Judicial Mind*, 86 CORN. L. REV. 777, 787-93 (2001) (reporting experimental findings on judicial cognition); Casey, Burke & Leben, *supra* note 63, at 81-82 (explaining cognitive heuristics and how they can both help and mislead in decision-making).

73. See Guthrie, Rachlinski, & Wistrich, *supra* note 72, at 787-88, 790, 793.

74. Jeffrey J. Rachlinski, Andrew J. Wistrich, & Chris Guthrie, *Can Judges Make Reliable Numeric Judgments? Distorted Damages and Skewed Sentences*, 90 IND. L.J. 695, 698-700 (2015) (arguing that rule-of-law values are threatened when numeric judgments vary widely across judges; emphasizing consistence as a legitimacy requirement); Englich, Mussweiler, & Strack, *supra* note 43, at 188-93 (demonstrating anchoring effects on sentencing recommendations).

75. Englich, Mussweiler, & Strack, *supra* note 43, at 192.

76. *Id.*

77. *Id.* at 190, 192-93. The high-anchor group had a mean probation of 6.05 months; the low-anchor group had a mean probation of 4.0 months. *Id.* at 192-93. There were thirty-seven judges and two prosecutors in the study group; in Germany, prosecutors and judges receive the same training. *Id.* at 190-91.

sentence even when the prosecutor's recommendation was obtained by the judge throwing a pair of loaded dice.⁷⁸

Many other cognitive biases and errors have been observed in judges. Among them:

Egocentric bias. Egocentric bias is the tendency to see our own skills and judgments as better than those of our peers—especially when the standard is vague and self-assessment is easy to rationalize.⁷⁹ A classic study of student drivers illustrates the point: ninety-three percent of the United States sample rated their driving skill as above average.⁸⁰ So too for judges. Chris Guthrie, Jeffrey Rachlinski, and Andrew Wistrich asked federal magistrates at a conference to estimate their appeal-reversal rate compared to other judges in the room.⁸¹ They were asked to put themselves into a quartile.⁸² Fifty-six percent put themselves into the least-reversed quartile, and eighty-eight percent put themselves into the top half.⁸³

Confirmation bias. Confirmation bias is the tendency to seek, interpret, and remember information in ways that support an initial hypothesis or belief while giving less weight to inconsistent evidence.⁸⁴ In ordinary life, once people form a view—about politics, a colleague, or a contested policy—they often evaluate supporting information as more credible and opposing information as more flawed, even when both sides are presented with the same mixed evidentiary support.⁸⁵ Over time, this can

78. Englich, Mussweiler, & Strack, *supra* note 43, at 195-96 (discussing implications of the robust effects of anchoring for judicial sentencing decisions).

79. See generally Wistrich, *supra* note 70, at 913-17 (explaining the general concept and research on judges while characterizing egocentric bias as “illusory superiority” and the “better-than-average effect”).

80. Ola Svenson, *Are We All Less Risky and More Skillful Than Our Fellow Drivers?*, 47 ACTA PSYCHOLOGICA 143, 146 (1981) (reporting that ninety-three percent of a United States student sample rated their driving skill above the median driver and eighty-eight percent rated their safety above the median).

81. Wistrich, *supra* note 70, at 913-14.

82. *Id.*

83. Guthrie, Rachlinski, & Wistrich, *supra* note 72, at 813-14 (discussing overconfidence in judicial assessments); see also Richard Schaffler & Kevin S. Burke, *Who Are You Going to Believe?*, 49 CT. REV. 124, 124-28 (213) (explaining that most people—including judges—perform poorly at lie detection; summarizing research that training does not reliably improve accuracy).

84. JONATHAN BARON, THINKING AND DECIDING 200 (5th ed., 2024).

85. *Id.*

produce irrational belief persistence as views that should become weaker in the face of mixed evidence instead become more confident.⁸⁶ For a judicial example, pretrial-detention decisions—made with very little information on hand—can trigger confirmation bias in later guilt assessments,⁸⁷ and there are many other judicial opportunities for early information exposure: preliminary credibility assessments, early exposure to case theories, and preliminary hearings in criminal cases (in which defendants rarely have an interest in exposing their case), to name a few.⁸⁸

Implicit bias. Implicit bias refers to automatic associations—learned mental linkages that may operate under the radar, without conscious endorsement.⁸⁹ Millions of people have taken the Implicit Association Test available online from Harvard’s Project Implicit.⁹⁰ Even people who consciously reject notions of prejudice often show faster reaction times pairing socially privileged groups with positive concepts.⁹¹ White judges taking those tests showed an overall white preference, just as the public does; black judges did not show a racial preference.⁹² Another study found that state and federal judges showed implicit bias

86. See *id.* at 204-06 (describing confirmation bias as “biased assimilation,” including how people credit supporting evidence and discount contrary evidence).

87. Moa Lidén, Minna Gräns, & Peter Juslin, “*Guilty, No Doubt*”: *Detention Provoking Confirmation Bias in Judges’ Guilt Assessments and Debiasing Techniques*, 25 PSYCH., CRIME & L. 219, 223, 238-40 (2019) (reporting a study of Swedish judges; detention and custody cues increasing confirmation bias).

88. *Id.* at 220-22 (showing detention increases guilt-confirming interpretations and testing debiasing techniques); BARON, *supra* note 84, at 204-06 (defining and explaining confirmation bias).

89. See generally Jeffrey J. Rachlinski et al., *Does Unconscious Racial Bias Affect Trial Judges?*, 84 NOTRE DAME L. REV. 1195, 1196-1201, 1225-26 (2009) (testing whether trial judges show implicit racial bias on IAT measures and whether it affects case-related judgments); Wistrich, *supra* note 69, at 903-07 (surveying judicial-cognition research and discussing implicit-bias mechanisms); Casey, Burke, & Leben, *supra* note 62, at 84-85 (describing implicit bias as an automatic schema-like process affecting judgment).

90. *Preliminary Information*, PROJECT IMPLICIT, [https://perma.cc/RSH7-ZV32] (last visited Mar. 11, 2026).

91. Casey, Burke, & Leben, *supra* note 63, at 83-84 (defining implicit biases as below-the-radar schemas and describing their potential to affect fair processes and outcomes).

92. Rachlinski et al., *supra* note 89, at 1210, 1225-26 (reporting results on judges’ implicit racial associations and discussing the relationship between implicit measures and judicial judgment patterns).

against Asians and Jews on the Implicit Association Test.⁹³ Wistrich, a former federal magistrate judge, concluded that “judicial expertise and experience do not [seem to] cure the tendency for most people to express implicit biases.”⁹⁴

Framing. Framing effects occur when judgment shifts because the same choice is presented with a different reference point (or “frame”).⁹⁵ For example, whether to settle civil litigation at a specific value might differ based on the plaintiff’s perspective (locking in some gain) or the defendant’s (agreeing to a known loss of money). Based on this framing of gaining or losing money, plaintiffs are generally more willing to settle a case.⁹⁶ Judges are subject to framing effects too, as shown in an experiment having half of a group of real judges evaluate a settlement value in a hypothetical case from the plaintiff’s perspective and half from the defendant’s perspective. Guthrie, Rachlinski, and Wistrich found a statistically significant difference in judicial evaluation of settlement value just based on framing.⁹⁷ As they note, judges often handle mediations and settlement conferences,⁹⁸ so differences like that would be significant in real cases too.

Hindsight bias. Hindsight bias is the tendency, once an outcome is known, to overstate how predictable it was beforehand.⁹⁹ In ordinary life, this can show up after accidents, stock-market swings, or medical outcomes: once we know what happened, we reconstruct the past as more predictable than it really was. If a sports team is behind with very little time left on the clock—but we know that this team won the game—we are

93. Justin D. Levinson, Mark W. Bennett, & Koichi Hioki, *Judging Implicit Bias: A National Empirical Study of Judicial Stereotypes*, 69 FLA. L. REV. 63, 104-10 (2017) (reporting results on stereotypes and associations among judges and the implications).

94. Wistrich, *supra* note 70, at 904 (discussing specific judicial-cognition evidence relevant to implicit bias).

95. Guthrie, Rachlinski, & Wistrich, *supra* note 72, at 794.

96. *Id.* at 794-96 (finding plaintiff-frame increased settlement recommendations relative to defendant frame).

97. *Id.* at 796-99 (reporting that forty percent of plaintiff-focused judges recommended settlement as compared to twenty-five percent of defendant-focused judges, with the result meeting statistical-significance tests).

98. *Id.* at 798 (explaining that judges often play an active role supervising settlement talks and “frequently serve as mediators during conferences,” so framing-driven differences in settlement valuation would matter in real cases).

99. *Id.* at 799.

likely in hindsight to overestimate what the winning chances really were. This can show up in a legal context in any situation in which past behavior is at issue. In child-neglect cases, for example, if a tragic outcome has already occurred, one state supreme court noted that a prosecution for neglect is especially vulnerable to “hindsight bias and risk distortion.”¹⁰⁰ Judicial experiments find judges subject to this bias too.

Significantly, despite evidence that judges are prone to the same cognitive errors non-judges are, our system depends on judges to be different. As one example, when trials occur, they are more often bench trials than jury trials, especially in civil cases.¹⁰¹ In those cases, judges rule both on the admissibility of evidence and the result. We rely on judges to ignore inadmissible evidence. But can they?

Research strongly suggests the answer is no.¹⁰² In controlled studies, judges’ decisions were systematically influenced by information they were told to disregard, including inadmissible settlement demands and other excluded materials.¹⁰³ While judges often discounted that information, they did not reliably eliminate its effect.¹⁰⁴ Even so, the same line of research reports that judges can sometimes resist some legally irrelevant

100. *State v. Cummings*, 305 P.3d 556, 565 (Kan. 2013). In *Cummings*, a day-care provider put a child into a car seat instead of a crib and fastened the top strap but not the bottom one. *Id.* at 558. The child was in the next room from the provider with the door open a few inches. *Id.* at 558-59. The child cried for a few minutes but then was quiet. *Id.* at 559. The caregiver assumed the child was asleep. *Id.* Unfortunately, the child had strangled to death. *Id.* The court noted that the result had skewed analysis, stating that “it is difficult to imagine that the State would have even prosecuted . . . in the absence of [the child’s] death, *i.e.*, the influence of hindsight bias was enormous.” *Cummings*, 305 P.3d at 567. *See also* *State v. Cummings*, 243 P.3d 697, 708 (Kan. Ct. App. 2010) (Leben, J., dissenting) (“[W]ith the hindsight bias we all have when we know how badly things turned out in a case like the one now before us, jurors can be tempted to interpret the words [“reasonable chance” of injury] liberally toward finding that a crime occurred.”), *rev’d*, 305 P.3d 556 (2013).

101. *See* BRIAN J. OSTROM & NEAL B. KAUDER, NAT’L CTR. FOR STATE CTS., EXAMINING THE WORK OF STATE COURTS, 1993: A NATIONAL PERSPECTIVE FROM THE COURT STATISTICS PROJECT 14 (1995) (reporting that, overall, 7.6% of civil filings were disposed by trial; 6.4% by bench trial and 1.2% by jury); *Judicial Business 2024: U.S. District Courts—Trials Completed*, U.S. CTS (2024), [<https://perma.cc/XW4F-Q96Z>] (reporting 2,133 civil non-jury trials versus 1,162 civil jury trials, and 5,416 criminal non-jury trials versus 1,581 criminal jury trials).

102. Wistrich, Guthrie, & Rachlinski, *supra* note 3, at 1278-79.

103. *Id.* at 1289, 1292, 1296, 1302, 1324.

104. *Id.* (reporting that trial judges struggled to disregard relevant but inadmissible information).

information.¹⁰⁵ The problem seems to be a cognitive processing failure, not bad faith. So, it's possible that some change in how judges process information could lead to better results.

Overall, though, decision science presents troubling questions about judges given the importance of the errors known to occur. This research is sometimes presented at judicial-education conferences, just as the work of procedural-justice researchers is presented. But an educational presentation or two is not likely to change daily judicial behavior months or years after the presentation is made.¹⁰⁶ It is hard to overcome noise, egocentric bias, confirmation bias, and other issues with any audience. Judges may be especially difficult because judicial error—and public reversal by a higher court—is often considered a personal failing rather than a learning experience. Intellectual humility, covered in the next Part of this Article, provides a motivational and social bridge that could turn self-improvement into a professional norm of craftsmanship. With intentional design, it can become a sustainable, institutionally normal process that gets built into normal court functions.

IV. INTELLECTUAL HUMILITY AND JUDICIAL TRAITS THAT SUPPORT LEGITIMACY AND RELIABILITY

If procedural justice helps explain why people accept and comply with court decisions, and decision science helps to explain why even conscientious decision-makers make avoidable errors and inconsistent rulings, intellectual humility provides a path to improve both legitimacy and reliability at the same time. Procedural justice has a direct relationship to popular legitimacy—an influence that flows most strongly through interpersonal treatment that leaves court participants feeling heard by judges who are trustworthy and concerned about the

105. *Id.* at 1322 (noting that judges generally could ignore an inadmissible confession in one scenario); Rachlinski et al., *supra* note 89, at 1221 (noting that judges were sometimes able to limit the effects of implicit bias when consciously aware of a need to monitor their responses for it).

106. KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 237-41 (noting that many one-time debiasing trainings show improved performance during the training exercise followed by rapid decay of the training effect).

public.¹⁰⁷ This effect is strong enough that an extensive California study found that perceptions of procedural justice were by far the strongest predictor of public confidence in the courts.¹⁰⁸

This legitimacy is earned by the way courts exercise their authority in public view. Fair procedures may well produce fairer outcomes. But fair procedures also communicate status and respect—messages that shape identification with legal authorities.¹⁰⁹ That interpersonal channel is where intellectual humility plausibly operates in courts to promote procedural justice: intellectual humility makes it more likely that judges will listen in ways that reduce defensiveness, convey genuine consideration, and treat contrary arguments as something to be engaged rather than brushed aside.

Decision reliability poses a parallel challenge. No one wants something important to them to be decided by chance. In a justice system, people expect consistent judgments in similar cases. But the noise that Kahneman, Sibony, and Sunstein describe can cause unwanted variability in which identical cases randomly assigned to different decision-makers may yield dramatically different results.¹¹⁰ And common cognitive errors could result in significant errors in case outcomes. Both the noise and cognitive-error problems are hard to solve. Accordingly, one core implication for Part V is that routines should be embedded into judging that make error-correction and processes to improve

107. See Tyler & Sevier, *supra* note 12, at 1097 (noting that the strongest influence of procedural justice on legitimacy flows through interpersonal treatment and perceived trustworthiness).

108. See Steve Leben, *Exploring the Overlap Between Procedural-Justice Principles and Emotion Regulation in the Courtroom*, 9 OÑATI SOCIO-LEGAL SERIES 852, 856 (2019) (noting the California study finding that perceptions of procedural justice were “the strongest predictor by far” of public confidence in the California state court system).

109. See Tyler & Sevier, *supra* note 12, at 1106 (explaining that quality of treatment is “a statement about status” and that people are sensitive to dignity, politeness, and whether rights are respected).

110. See Jaya Ramji-Nogales, Andrew I. Schoenholtz & Phillip G. Schrag, *Refugee Roulette: Disparities in Asylum Adjudication*, 60 STAN. L. REV. 295, 339 (2007) (stating that asylum outcomes depend in large part on the assigned judge, describing within-building grant rates ranging from five percent to eighty-eight percent in Miami); KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 43, at 6-8 (stating that asylum decisions are noisy and that whether an asylum seeker is admitted “depends on something like a lottery[,]” with judge-to-judge admission rates ranging widely).

consistency structured and professionally acceptable.¹¹¹ Those are roles that intellectual-humility research can help with.

Intellectual humility is defined by social psychologists as “a metacognitive core composed of recognizing the limits of one’s knowledge and awareness of one’s fallibility.”¹¹² It is an awareness of fallibility and the limits of evidence that supports both truth-seeking and revision of one’s views when warranted.¹¹³ Validated tests have been developed to measure intellectual humility under this definition,¹¹⁴ which looks only at mental or cognitive processes. Some social psychologists broaden the definition to include social and behavioral aspects, like valuing other people’s beliefs or how one’s behavior shows intellectual humility.¹¹⁵ Validated measures have also been developed for that.¹¹⁶

With these definitions, intellectual humility is not the same as general open-mindedness or modesty. As studied in social psychology research, intellectual humility is chiefly concerned with epistemic limitations and fallibility.¹¹⁷ It reflects a disciplined, calibrated confidence that allows decision-making with an active awareness that one’s understanding may be partial

111. See KAHNEMAN, SIBONY, SUNSTEIN, *supra* note 44, at 21-22 (defining noise as “unwanted variability” and emphasizing that noise can exist even when decisions seem unique); Casey, Burke, & Leben, *supra* note 63, at 76-77 (framing why judges need decision-making routines and decision aids or checklists to reduce error and improve consistency); *see infra* Part V.

112. Tenelle Porter et al., *Predictors and Consequences of Intellectual Humility*, 1 NATURE REV. PSYCH. 524, 525 (2022).

113. *Id.*

114. *Id.* at 525-28.

115. *Id.* at 525-26 (distinguishing metacognitive “core” elements from social-behavioral components and discussing broader accounts that include valuing others’ beliefs and expressed behavior); *see also* Elizabeth J. Krumrei-Mancuso & Steven V. Rouse, *The Development and Validation of the Comprehensive Intellectual Humility Scale*, 98 J. PERSONALITY ASSESSMENT 209, 209-10 (2016) (defining intellectual humility as including intrapersonal and interpersonal elements); Tenelle Porter & Karina Schumann, *Intellectual Humility and Openness to the Opposing View*, 17 SELF & IDENTITY 139, 140 (2018) (including appreciation of others’ intellectual strengths in the operationalization of intellectual humility).

116. *See* Porter et al., *supra* note 112, at 525-27 (describing multidimensional accounts of intellectual humility and summarizing multiple measurement approaches); Krumrei-Mancuso & Rouse, *supra* note 114, at 212-14, 216 (developing and validating a twenty-two-item scale measuring facets including respect for others’ viewpoints and openness to revising one’s viewpoint).

117. Porter et al., *supra* note 112, at 524-25.

and that strong counterarguments and contrary evidence must be addressed, not discounted or ignored.¹¹⁸

A key reason intellectual humility matters for judges is that it is often hardest to sustain when it is most needed. People are less likely to exhibit intellectual humility when the stakes are high, especially when strong convictions or identity-relevant values feel challenged. Yet judges work on high-stakes disputes, with adversarial pressures, on a stage (well, usually a raised judicial bench) in public. It also explains why measurement matters. Researchers describe intellectual humility as measurable both as a relatively stable disposition and as a context-specific response, and they have developed measures for both.¹¹⁹

Management expert Peter Drucker said that “[w]ork implies accountability . . . [and] the measurement of results,” and that “[w]hat we measure and how we measure determine . . . not just what we see, but what we—and others—do.”¹²⁰ Measures of intellectual humility can help courts and judges in similar ways. They can function as a mirror for judicial decision-making by showing blind spots that are easy to overlook in high-stakes, time-pressured adjudication—like overconfidence, premature closure, or the tendency to discount the strongest counterargument once a tentative view has been formed. Even when a scale is inward-facing, it can identify discrete dimensions, like openness to revising one’s viewpoint or respect for others’ viewpoints, that can be translated into observable decision-making habits (for example, how a judge frames questions, acknowledges counterarguments, and explains reasons). Measurement also makes reform claims testable by providing baselines and outcome measures for evaluating training and design changes over time. With that in mind, we turn to what the leading measures of intellectual humility capture; they supply a practical blueprint for the judicial behaviors and institutional scaffolds set out in this Article. Used this way, these measures can help define and reinforce professional routines that make careful judging more reliable and more visible to court users.

118. *Id.* at 525.

119. *Id.*

120. Joe Maciariello, *Joe’s Journal: On Measuring Performance*, DRUCKER INST. (Aug. 30, 2011), [<https://perma.cc/GF4L-36CN>] (quoting Peter F. Drucker).

Measurements have been designed that separate intellectual humility as a concept from broader, normative accounts of virtue. The Comprehensive Intellectual Humility Scale is a twenty-two-item instrument developed through expert review, pilot testing, and factor analysis.¹²¹ It measures four distinct but interconnected concepts: independence of intellect and ego, openness to revising one's viewpoint, respect for others' viewpoints, and lack of intellectual overconfidence.¹²² Those dimensions map naturally onto the behaviors people watch for when deciding whether a judge is neutral and open to reason. Even so, judges' self-report measures can only be treated as indicators, not guarantees: intellectual humility is context-sensitive, and what matters for courts is how it is expressed under the pressures of everyday judging, not how it is endorsed in the abstract.¹²³

Researchers have reported meaningful divergence between self-reported intellectual humility scores and behavior. In one study, trained coders scored intellectual humility from participants' debate discourse, and researchers then tested an intervention to reduce defensiveness.¹²⁴ The study found that a brief self-affirmation task increased coded intellectual humility, while self-reported intellectual humility and other personality measures were not correlated with the expert-coded measures.¹²⁵ The lesson here is although the validated intellectual-humility scales are valuable, they are incomplete. We must also be concerned professionally about how intellectual humility is expressed under the pressures inherent in judging, and that makes field-facing practices and institutional supports especially important. That behavioral emphasis also aligns with what research consistently finds as a downstream consequence of

121. Porter et al., *supra* note 112, at 212-14.

122. Krumrei-Mancuso & Rouse, *supra* note 114, at 215 (describing the twenty-two-item Comprehensive Intellectual Humility Scale and its four subscales); *see also* Mark K. Leary et al., *Cognitive and Interpersonal Features of Intellectual Humility*, 43 PERSONALITY & SOC. PSYCH. BULL. 793, 794-95 (2017).

123. Porter et al., *supra* note 112, at 525.

124. Paul H. P. Hanel et al., *Using Self-Affirmation to Increase Intellectual Humility in Debate*, ROYAL SOC'Y OPEN SCI., Feb. 1, 2023, at 1, 2-3, 10.

125. *Id.* at 3, 10 (reporting that self-affirmation increased behaviorally coded intellectual humility in debate and that trait self-reports were essentially uncorrelated with the coded measure); *see also* Porter et al., *supra* note 112, at 526. (discussing limits of trait questionnaires and the value of situational or contextualized assessment).

intellectual humility—greater openness to opposing views and less myside bias.¹²⁶

Of course, greater openness to the views of others is important from a procedural-justice viewpoint because it supports the legitimacy signals people care about: voice, neutrality, and respectful treatment. Those with greater intellectual humility are consistently associated with greater respect for and openness to opposing views and with greater willingness to seek out and engage reasons they expect to reject.¹²⁷ Courts earn trust in the same way: by demonstrating open, fair engagement with the strongest arguments before deciding against them, and by explaining why those arguments do not carry the day.

Political psychology literature offers a useful yet troubling analogue for courts confronting identity-salient disputes: motivated reasoning often produces myside bias, the tendency to evaluate evidence and arguments in ways that favor one's side.¹²⁸ In a large study of political reasoning, measures of intellectual humility were negatively correlated with political myside bias across paradigms, topics, and samples, and those relationships were statistically equivalent across people on the political left and right.¹²⁹ In a time when judges routinely decide disputes that track sharp partisan divides,¹³⁰ myside bias poses both a decision-making risk and a legitimacy risk.¹³¹ Even if a particular judge is not operating under myside bias, litigants and the public may reasonably suspect that judges will favor "their side" in identity-salient cases.

126. Leary, *supra* note 4, at 10.

127. See generally Krumrei-Mancuso & Rouse, *supra* note 115, at 209, 211-12 (linking intellectual humility with openness to learning about opposing views and willingness to engage opposing perspectives); see also Tyler & Sevier, *supra* note 12, at 1100-07 (connecting respectful treatment and voice to perceived status and legitimacy signals).

128. Bowes et al., *supra* note 5, at 150.

129. *Id.* at 152.

130. See *Partisan Judges*, *supra* note 14, at 520-21 (cataloging sharp partisan divides on major policy issues that routinely generate litigation—including abortion, guns, immigration, environmental regulation, and agency power).

131. See Bowes et al., *supra* note 5, at 151-52 (reporting negative correlation between intellectual humility and political myside bias across paradigms and symmetry across left and right political views); see also Porter et al., *supra* note 112, at 531 (noting that intellectual humility correlates with greater intergroup tolerance).

Even so, no single trait can overcome motivated reasoning, and the literature on actively open-minded thinking illustrates both the promise and the limits of trait-based approaches. Intellectual humility and actively open-minded thinking are related constructs in the measurement of open- versus closed-mindedness.¹³² They overlap most clearly in the belief-revision space: both value attention to contrary arguments and evidence, delaying premature closure, and willingness to revise opinions when warranted. But they are not fully identical. Intellectual humility centers on metacognitive awareness that one's knowledge may be limited and one's beliefs fallible. Actively open-minded thinking comes from a rational-thought tradition and focuses more directly on whether a person is disposed to consider alternatives, search for reasons against a favored view, and revise beliefs in light of evidence.¹³³ Despite these differences, considering both could be useful here. One might prefer the selection of judges with greater intellectual humility or work to train current judges to have more of it, assuming trainability. And one might practice actively open-minded thinking as part of judicial routines.

Empirically, though, research has shown some differences. Intellectual-humility measures have been negatively associated with political myside bias across paradigms, topics, and samples.¹³⁴ By contrast, actively open-minded thinking scales—despite predicting many other reasoning outcomes—have shown little or no relationship with direct measures of myside thinking, a result some researchers have described as a serious problem for

132. See Leary, *supra* note 4, at 10 (situating intellectual humility among related open- and closed-mindedness constructs); Keith E. Stanovich & Maggie E. Toplak, *Actively Open-Minded Thinking and Its Measurement*, J. INTEL., JAN. 28, 2023, at 1, 2, 10-12 (describing actively open-minded thinking items and its measurement); Krumrei-Mancuso & Rouse, *supra* note 114, at 215 (including “openness to revising one’s viewpoint” as a component of intellectual humility).

133. See Porter et al., *supra* note 112, at 132 (defining intellectual humility’s metacognitive core as recognizing the limits of one’s knowledge and awareness of one’s fallibility); Stanovich & Topic, *supra* note 132, at 2, 10-12 (describing actively open-minded thinking items involving consideration of alternatives, postponement of closure, and revision of beliefs in light of evidence).

134. See Bowes, *supra* note 115, at 150, 156-62 (reporting negative associations between intellectual humility and political myside bias and finding little evidence of political asymmetry in those relationships).

that construct.¹³⁵ And in one study of polarization over climate change, higher actively open-minded thinking scores were associated with *more* divergence between liberals and conservatives.¹³⁶

The point here is that neither actively open-minded thinking nor intellectual humility can be relied on as self-executing personality traits. They bring promise, based on evidence, for improving fairness, consideration of opposing views, and weighing all the evidence—all valuable things for judges to do. But courts will need structural scaffolds—like premortems, second looks, counterargument assignments, and reasons-first explanation—that make humility-relevant behavior more likely when adversarial pressure, public visibility, and identity-salient disputes make motivated reasoning the most likely. Meanwhile, caution against relying on any single trait is good reason to try to weave together multiple strands all designed to move toward better procedural justice, court legitimacy, and decision-making.

Two intervention paths seem especially relevant for a “by design” approach. First, self-affirmation, often used to reduce defensiveness, can increase observed intellectual humility in adversarial debate settings.¹³⁷ Second, perspective tools that create psychological distance can improve wise reasoning about one’s own conflicts.¹³⁸ In experiments on “Solomon’s paradox,” the tendency to reason more wisely about other people’s problems than our own, instructing individuals to self-distance eliminated the self-versus-other asymmetry in wise reasoning.¹³⁹ While these

135. See Stanovich & Toplak, *supra* note 132, at 12-13 (discussing the failure of actively open-minded thinking scales to predict myside bias despite their relationship to many other reasoning outcomes).

136. See Dan M. Kahan & Jonathan C. Corbin, *A Note on the Perverse Effects of Actively Open-Minded Thinking on Climate-Change Polarization*, RSCH. & POL. Oct.-Dec. 2016, at 1, 2-3 (noting that polarization increases as actively open-minded thinking scores increase, providing predicted probability differences).

137. Hanel et al., *supra* note 124, at 3, 8, 10-12 (reporting that a self-affirmation intervention previously used to decrease psychological defensiveness increased behaviorally coded intellectual humility in debates).

138. Igor Grossmann & Ethan Kross, *Exploring Solomon’s Paradox: Self-Distancing Eliminates the Self-Other Asymmetry in Wise Reasoning About Close Relationships in Younger and Older Adults*, 25 PSYCH. SCI. 1571, 1571-72 (2014) (defining Solomon’s paradox and introducing self-distancing to overcome it).

139. *Id.* at 1575-76 (reporting that instructing individuals to self-distance eliminated the self-other asymmetry in wise reasoning).

studies do not show that judges can be trained into stable-trait intellectual humility, they are promising evidence-based levers.¹⁴⁰ Procedures that reduce defensiveness and create distance from initial intuitions might make humility-relevant behaviors more likely in exactly the moments when adversarial presentation would otherwise trigger rigid, self-protective reasoning.

Among humility-adjacent practices, high-quality listening has unusually strong evidence because it is behaviorally specified and experimentally tested.¹⁴¹ In work on disagreements, high-quality listening—focusing on attention, accurate understanding, and a nonjudgmental intent—is thought to increase self-insight and reduce attitude extremity in ways that support reflective reconsideration.¹⁴² An effective listener aims to understand the message the speaker is actually trying to convey, and simple techniques can help a judge keep focused during a hearing.¹⁴³ One of the most important is confirming understanding: the listener briefly summarizes what was heard and asks whether that summary is right.¹⁴⁴ Done well, that supports both perceived fairness (people feel heard and respected) and better judgment

140. *Id.* at 1578 (suggesting that self-distancing may provide a tool to enhance wise reasoning when people have personal stakes); Hanel et al., *supra* note 124, at 10 (framing self-affirmation as a defensiveness-reducing intervention in contentious discussion); see Jack B. Soll, Katherine L. Milkman & John W. Payne, *A User's Guide to Debiasing*, in *THE WILEY BLACKWELL HANDBOOK OF JUDGMENT AND DECISION MAKING* 930-40 (Gideon Keren & George Wu eds., 2015) (recommending debiasing strategies that are built into decision procedures and targeted to the decision context).

141. Guy Itzhakov et al., *Listening to Understand: The Role of High-Quality Listening on Speakers' Attitude Depolarization During Disagreements*, 126 J. PERSONALITY & SOC. PSYCH.: ATTITUDES & SOC. COGNITION 213, 217 (2024) (describing high-quality listening as varying on three “core qualities” of listening—attention, comprehension, and positive intention—and operationalizing those qualities through scripted verbal and nonverbal behaviors).

142. *Id.* at 218-19.

143. *Id.* at 214-15; see *Expectation of Empathy*, *supra* note 23, at 54 (noting the use of training on listening skills in judicial education programs).

144. Itzhakov et al., *supra* note 140, at 218-19 (giving an example of confirmatory reflection—“I would like to make sure that I understood you correctly”—as part of the high-quality listening condition and describing a listening protocol, including the instruction to provide summaries beginning, “I want to make sure I understand . . .”).

(the judge is less likely to decide based on a misunderstanding of the party's point).¹⁴⁵

Because judging is often collegial—appellate panels, court conferences, chambers consultations, committees—intellectual humility in the judiciary should be understood not only as an individual attribute but also as an institutional one that can be supported or undermined by design features.¹⁴⁶ Judges, like other professionals, often respond to norms and incentives that reward solidarity and discourage open conflict.¹⁴⁷ Researchers studying intellectual humility therefore suggest that groups (which they call collectives) can display intellectual humility when they recognize and attend both to members' limitations and to the limits of the group's own reasoning processes.¹⁴⁸ They emphasize that collective humility is not simply an average of individual humility but is shaped by norms, roles, and procedures.¹⁴⁹ That insight is especially important for appellate courts because structural pressures can suppress disagreement: dissent aversion and workload can discourage dissenting even when a judge disagrees, and dissent rates are empirically related to caseload and ideological diversity.¹⁵⁰ The implication is that courts should treat constructive disagreement and find potential errors before decisions are made as design problems—requiring explicit

145. See Tyler & Sevier, *supra* note 12, at 1129-30 (explaining that perceived dignity, respect, and trust in the motives of legal authorities are central to the interpersonal side of procedural justice and legitimacy).

146. Krumrei-Mancuso et al., *supra* note 5, at 23 (arguing that institutions can cultivate collective intellectual humility through norms and “environmental scaffolding” rather than relying solely on individual traits).

147. See Porter et al., *supra* note 112, at 529 (explaining that group conformity and status motives can work against intellectual humility); Lee Epstein, William M. Landes & Richard A. Posner, *Why (and When) Judges Dissent: A Theoretical and Empirical Analysis* 4-6 (John M. Olin L. & Econ. Working Paper No. 510, 2010) (explaining that dissenting can impose “collegiality costs” and that future withholding of collegiality can deter dissent).

148. See Krumrei-Mancuso et al., *supra* note 5, at 17 (defining collective intellectual humility as the tendency of collectives to recognize and attend to members' and group-level intellectual limitations).

149. *Id.* at 19, 21.

150. Lee Epstein, William M. Landes, & Richard A. Posner, *Why (and When) Judges Dissent: A Theoretical and Empirical Analysis*, 3 J. LEGAL ANALYSIS 101, 102-03 (2011) [hereinafter *Why (and When) Judges Dissent*] (noting both dissent aversion and that dissent is often negatively related to caseload and positively related to ideological diversity).

scaffolds that legitimize doubt, invite counterarguments, and prevent premature consensus.¹⁵¹

That design challenge is ultimately a leadership task. Appellate courts and court systems will be more consistently intellectually humble only when chief judges and colleagues build norms, incentives, and routines that make constructive doubt and disagreement safe and expected.¹⁵² So, a focus on court leadership is needed too. In organizational research, “leader expressed humility” is defined as an interpersonal characteristic that shows up in social interactions through three behaviors: a willingness to view oneself accurately, an appreciation of others’ strengths and contributions, and teachability (openness to new ideas, feedback, and advice).¹⁵³ A review of this literature links humble leader behavior to outcomes that courts care about—learning, grounded decision-making, and psychologically safer climates for raising concerns.¹⁵⁴ For chief judges and court administrators, that literature provides an evidence-based way to talk about humility not as a softness, but as a practical leadership resource that improves information flow and institutional learning.¹⁵⁵

151. See Krumrei-Mancuso et al., *supra* note 5, at 17, 21-22 (describing “scaffolds” and explicit roles and procedures that make it easier to surface limits and errors); Mary Burns & Rebecca Winthrop, *Imagining failure to attain success: The art and science of pre-mortems*, BROOKINGS (Feb. 6, 2025), [<https://perma.cc/MX6W-C4ED>] (explaining that pre-mortems help counter optimism bias and make it safer to surface adverse possibilities before committing).

152. See Krumrei-Mancuso et al., *supra* note 5, at 18-21 (explaining that collective intellectual humility depends on implementation levers such as leadership behaviors, incentives, procedures, and scaffolding); Casey, Burke, & Leben, *supra* note 63, at 96 (urging feedback loops and accountability practices that make better judging the default rather than an exception).

153. Anthony Silard et al., *Leader Expressed Humility: A Meta-Analysis and an Agenda for Future Research*, PERSONALITY & INDIVIDUAL DIFFERENCES, Aug. 2025, at 1, 1-2 (defining leader humility as an interpersonal characteristic consisting of an accurate self-view, appreciation of others, and teachability).

154. *Id.* at 3 (noting that leader-expressed humility is “co-created in social and relational interactions” and supports a developmental path as legitimate); see also Kevin S. Burke & Steve Leben, *Procedural Fairness in a Pandemic: It’s Still Critical to Public Trust*, 68 Drake L. Rev. 685, 701-04 (2020) (encouraging courts to solicit feedback and revise procedures when they are not working, particularly under stress) [hereinafter *Procedural Fairness in a Pandemic*].

155. See DANIELLE KALIL, INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., JPE 2.0 RECOMMENDATIONS FOR MODERNIZING JUDICIAL PERFORMANCE EVALUATION 23-24 (2025) (emphasizing evaluation systems that focus on concrete judicial behaviors like listening and explaining and that earn trust through transparency and reliable methods).

One practical way leaders can make that culture real is to adopt routines that institutionalize candor—formal moments when raising doubts and surfacing risks is expected, not penalized.¹⁵⁶ Part V develops those routines in detail.¹⁵⁷ As a preview, though, one scaffold that fits both collective intellectual humility and decision-hygiene goals is the premortem.¹⁵⁸ A premortem is a structured, forward-looking thought experiment in which a team assumes a project or decision has failed and asks why, then works backward to surface risks and mitigation steps before committing.¹⁵⁹ The premortem framework is explicitly designed to counter overconfidence and social-desirability pressure by making it psychologically safe to voice doubts and even out-of-the-box concerns.¹⁶⁰ Collective intellectual-humility frameworks treat premortems and second-look policies as examples of the scaffolds that enable intellectual humility at scale.¹⁶¹ Part V uses these insights to translate intellectual humility into repeatable routines for trial courts, appellate courts, and court administration.¹⁶²

156. See Silard et al., *supra* note 153, at 3 (defining leader-expressed humility and linking it to psychological safety and climates in which people feel safe raising concerns); Krumrei-Mancuso et al., *supra* note 5, at 21 (identifying leadership behaviors, incentives, and procedures as “implementation levers” for collective humility).

157. See *infra* Part V.

158. See KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 370-74 (introducing “decision hygiene” and explaining that preventive, ex ante interventions operate by changing the process before judgment).

159. Burns & Winthrop, *supra* note 151 (defining a premortem as a “forward-looking thought experiment” in which a team assumes failure and then works backward to identify risks and actions).

160. *Id.* (explaining that premortems reduce overconfidence and mitigate social-desirability bias by “creat[ing] safety for honest dialogue” and encouraging candid, “out-of-the-box” concerns).

161. Krumrei-Mancuso et al., *supra* note 5, at 23 (listing premortems and second-read policies as structural features that function as scaffolds enable humility at scale).

162. See KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 370 (analogizing decision hygiene to “handwashing” and describing it as a set of preventive procedures designed to reduce error before it occurs); see also Soll, Milkman, & Payne, *supra* note 139, at 926 (cataloging debiasing strategies and emphasizing the value of embedding debiasing into decision procedures rather than relying on exhortation).

V. INTELLECTUAL HUMILITY BY DESIGN

This Article's central claim is that three literatures—procedural justice, decision science, and intellectual humility—can be woven into a single framework for judicial work that is more powerful than any strand standing alone. This section puts that argument in place and in context. In doing so, the Article introduces the term Intellectual Humility by Design to describe the intentional use of intellectual-humility research to structure judicial work.

A. The Integration Model

Procedural justice explains the legitimacy mechanism: court users accept decisions, even adverse ones, when they experience fair process.¹⁶³ Decision science explains the reliability problem: judges, like all decision-makers, face predictable vulnerabilities to overconfidence, noise, framing effects, attentional limits, and motivated reasoning, especially under time pressure and often under emotional pressures as well.¹⁶⁴ Intellectual humility provides a strong connection for both. It provides a measurable and trainable posture toward reducing error and considering opposing views, and those features should help courts to adopt, sustain, and normalize error-checking safeguards and fair procedures.¹⁶⁵

The integration works both internally and externally.¹⁶⁶ Methods suggested by intellectual-humility research like second

163. See Tyler & Sevier, *supra* note 12, at 1096-97 (reporting the direct effect of procedural justice on court legitimacy); *Procedural Justice and the Courts*, *supra* note 17, at 30-31 (explaining the elements of procedural justice).

164. See generally Chris Guthrie, Jeffery Rachlinski, & Andrew Wistrich, *Blinking on the Bench: How Judges Decide Cases*, 93 CORN. L. REV. 1, 43 (2007) (reviewing cognitive errors commonly made by judges and suggesting methods to improve judicial decision-making) [hereinafter *Blinking on the Bench*]; KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 298-99 (explaining that noise is difficult to see without a statistical view and suggesting decision hygiene as a preventative).

165. See generally Porter et al., *supra* note 112, at 524 (reviewing research on intellectual humility and finding that it serves as a safeguard against errors and biases).

166. See *Blinking on the Bench*, *supra* note 163, at 29-30 (describing institutional reforms aimed at inducing more deliberative decision-making); KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 298-301 (describing decision hygiene measures); Krumrei-

looks, premortems, and structured engagement with counterarguments improve decision-making.¹⁶⁷ But they can also improve external perceptions of fairness when courts recognize and openly engage with the losing party's best arguments.¹⁶⁸

The result of designing intellectual humility into the process can be reduced defensiveness, which can make error-checking feel like a threat to the judge's authority, while supporting a group norm of openness to opposing ideas and to allowing disagreement to function as a source of information rather than a source of personal conflict.¹⁶⁹ Intellectual humility can be a professional method for individual judges, appellate panels, and court organizations.¹⁷⁰

The rest of this Part operationalizes intellectual humility by design with specific application in four settings: (1) trial-court practices that combine voice and respect with decision discipline;¹⁷¹ (2) collective practices for appellate panels and court administration that structure disagreement, review, and learning;¹⁷² (3) reasons and explanations that make intellectual humility observable to decision recipients;¹⁷³ and (4) design choices that help courts turn these practices into durable institutional routines rather than one-off aspirations.¹⁷⁴

Mancuso et al., *supra* note 5, at 22 (describing premortems and second-read policies as scaffolds enabling intellectual humility at scale).

167. See generally Porter et al., *supra* note 112, at 531 (noting that intellectual humility is positively associated with knowledge acquisition, a better ability to distinguish weak from strong arguments, and more scrutiny of misinformation).

168. See generally Tyler & Sevier, *supra* note 12, at 1097 (noting that when people feel that authorities "consider and take account of their concerns," acceptance of legal authorities is enhanced); Steve Leben, *Getting It Right Isn't Enough: The Appellate Court's Role in Procedural Justice*, 69 KAN. L. REV. 13, 32-34 (2020) (noting sense of procedural justice when parties "feel that the court listened to both sides") [hereinafter *Getting It Right Isn't Enough*]; see also BRYAN GARNER, THE REDBOOK A MANUAL ON LEGAL STYLE § 26.1(a) (5th ed. 2023) (stating that judicial opinions should justify outcomes "especially to the losing litigant" and not slight losing-side merits).

169. Porter et al., *supra* note 112, at 524.

170. See Krumrei-Mancuso et al., *supra* note 5, at 17 (explaining that "collective intellectual humility" applies to "group agents," including "teams, corporations, and juries," and is shaped by group structure, culture, and institutional scaffolding); Porter et al., *supra* note 112, at 530-31 (summarizing intellectual-humility research at the individual level)

171. See *infra* Section V.B.

172. See *infra* Section V.D.

173. See *infra* Section V.C.

174. See GARNER, *supra* note 168, at § 26.1(a) (describing judicial opinions as helping the judge think and justifying outcomes, especially for losing parties); Bridget Mary

B. Trial Court Practices

Trial judges operate in real time handling busy dockets with parties, attorneys, witnesses, and other observers at hand. Like a tennis player in a five-set match, a momentary loss of concentration can have consequences. But rather than losing a point, a set, or a match, the judge may make a substantive error that changes a case outcome. Or the judge may lead participants to believe that their case was not taken seriously. Building in practices that help judges maintain focus, make procedural justice observable, and improve judicial reasoning under time pressure can help keep judicial performance high despite the pressures of the trial-court docket.¹⁷⁵

Each practice in the intellectual-humility-by-design framework seeks to strengthen the information the court receives (voice and listening as inputs), discipline the court's reasoning (structured reflection and counterargument engagement), or clarify the path of decision (providing reasons that show method and constraint rather than mere outcome). The design embeds professional methods into routine practice.¹⁷⁶

Let's start with the inputs. People are more likely to accept what a judge decides when they feel heard, and voice is a core component of procedural justice.¹⁷⁷ Trial-court fairness studies repeatedly show that court users' satisfaction is driven by how they are treated—visible signs like being listened to and having participants understand what is expected—rather than by whether they win.¹⁷⁸ The practical point is that priority be given to making

McCormack, *Access to Justice Requires Plain Language*, MICH. BAR J., Feb. 2021, at 44, 44 (stating that the justice system is “not fair when litigants cannot understand it” and linking understandable explanations to fairness and legitimacy); and *See* Krumrei-Mancuso et al., *supra* note 5, at 23 (describing implementation layers and scaffolding as ways to make intellectual humility durable at scale).

175. *See* Casey, Burke, & Leben, *supra* note 63, at 76-77 (describing cognitive limits on judging and recommending practical routines—breaks, checklists, and feedback—to improve decision-making).

176. *See id.* at 91-96 (recommending decision aids, breaks, and mindfulness practices).

177. *Public Satisfaction*, *supra* note 23, at 5-6.

178. *See, e.g.*, GINA KUBITS ET AL., FOURTH JUD. DIST. OF THE STATE OF MINN., HOUSING COURT FAIRNESS STUDY 27-28 (2004) (reporting that outcome favorability—whether they won or lost—explained about seventeen percent of variance in satisfaction, while fairness, listening, and understanding explained another forty-three percent of variance).

at least some space for the party's story. In a busy trial docket, voice is often most workable as a short, structured window—enough to surface what matters, but bounded enough to keep the hearing manageable.¹⁷⁹

Listening must be visible. Court users don't infer careful consideration from a judge's inner state; they infer it from what they can observe in the courtroom. A Milwaukee courtroom-communication experiment is a useful guide because it treats procedural-justice practices as observable behaviors and measures them directly—eye contact, direct conversation, plain-English explanation of procedure and outcome, asking whether the criminal defendant has anything else to say before the decision, and checking understanding of next steps.¹⁸⁰ Those are not empty gestures. They reduce misunderstanding, signal that the judge is engaged, and can improve accuracy by making it more likely that the judge hears what is unique about this case rather than defaulting to the last similar one, even unconsciously.¹⁸¹

A short opening script is one of the simplest ways to institutionalize respect and predictability. In Milwaukee, judges made significant changes after training: from never explaining the order of cases to doing so in a majority of sessions, and from assuring that evidence would be considered in just over half of sessions to doing so in almost all sessions.¹⁸² Those moves lower stress, set expectations, and give court users a basic mental model for what will happen next.¹⁸³ They also address a point that shows up repeatedly in trial-court fairness studies: when people feel they

179. *Id.* (noting procedural-justice factors and finding satisfaction driven by fairness, listening, and understanding).

180. See ERIN J. FARLEY, ELISE JENSEN, & MICHAEL REMPEL, CTR. FOR CT INNOVATION, IMPROVING COURTROOM COMMUNICATION: A PROCEDURAL JUSTICE EXPERIMENT IN MILWAUKEE iv, 14-15 (2014) (listing measured courtroom behaviors including eye contact, direct conversation, plain-English explanation, asking for anything else before decision, and checking understanding/next steps).

181. See Casey, Burke, & Leben, *supra* note 63, at 77 (explaining that a judge focused on procedural justice will listen carefully and repeat key themes, creating the opportunity to notice how a case differs from previous ones); See Tyler & Sevier, *supra* note 12, at 1107 (noting that studies have found that allowing voice to the parties helps lead to appropriate factual findings by the court).

182. See FARLEY, JENSEN, & REMPEL, *supra* note 180, at 24-25.

183. *Id.* at iv, 14-15, 66 (describing the observation protocols and listing measured communication behaviors targeted by the intervention, including explaining call order and assuring defendants that all evidence would be considered).

waited too long or did not understand the process, they rate fairness and satisfaction lower even when outcomes are legally correct.¹⁸⁴

Plain language and participant understanding are part of the same package. Procedural justice is not only about allowing someone to speak; it also includes whether people understand what is happening and what is required next.¹⁸⁵ The Milwaukee protocol explicitly measured whether judges used plain English to explain procedure and outcome and whether they adequately described what the defendant must do to comply.¹⁸⁶ Hennepin County's Housing Court fairness study likewise treated "understand[ing] what is required . . . to comply" as a procedural-justice indicator and found that fairness, listening, and understanding together explained far more variance in satisfaction than a favorable outcome did.¹⁸⁷

A simple comprehension check—brief, respectful, and not performative—does double work: it improves decision quality and strengthens perceived fairness. One low-friction method is to ask for a summary of next steps ("Tell me what you understand you must do next") rather than asking only, "Do you understand?"¹⁸⁸ The Milwaukee study shows that courts can observe and measure whether judges ask about understanding and whether they use repeat-back techniques.¹⁸⁹ Even when a particular technique fluctuates after training, treating comprehension as a design target shifts the court's emphasis from

184. See KUBITS ET AL., *supra* note 178, at 27-28 (noting that perceived delay and confusion about procedures eroded fairness ratings in a busy court handling housing cases).

185. *Id.* (using "understand[ing] what is required . . . to comply" as a procedural-justice indicator) (italics removed).

186. See FARLEY, JENSEN, & REMPEL, *supra* note 180, at 24, 26, 69 (measuring "used colloquial English to explain case procedure and outcome" and "adequately described what defendant must do to comply").

187. See KUBITS ET AL., *supra* note 178, at 27-28 (measuring "understand[ing] what is required . . . to comply") (italics removed).

188. See FARLEY, JENSEN, & REMPEL, *supra* note 180, at 14-15, 70 (measuring whether the judge asked if the defendant understood next steps).

189. *Id.* at 26 (measuring whether the judge asked if the defendant understood next steps and whether the judge asked defendant to repeat back understanding of the next steps).

delivering an outcome to delivering an outcome that can actually be followed.¹⁹⁰

Explanations matter most when the judge says “no.” One of the strongest trial-court findings is that receiving an explanation from the bench made the biggest difference in satisfaction for litigants who had a full trial—and the effect is strongest for trial litigants who did not get what they asked for.¹⁹¹ Even in highly contested domestic-violence cases, one study reported that trial losers were more likely to say they would comply when they reported both fair treatment and having received a full explanation.¹⁹² Reason-giving, along with an opportunity to ask questions, can change how litigants experience even adverse outcomes.¹⁹³

A brief question-and-answer opportunity at the end of a ruling is an underused trial-court tool. In the Hennepin domestic-abuse study, the “explain” condition required stating the outcome and reasons and then inviting questions.¹⁹⁴ That design is valuable in two respects. It is a procedural-justice intervention—showing that the court took the matter seriously and addressed what mattered—and it is a decision-quality intervention because questions expose misunderstandings immediately, when they can still be corrected before the court’s ruling becomes a miscommunicated directive.¹⁹⁵

None of this requires any fundamental change in the judge’s role, but it can provide a professional discipline of attention and method under time pressure. Decision-making research aimed

190. *Id.* at iv (treating comprehension-related behaviors as observable and reportable, supporting follow-through as a design problem).

191. DEBORAH A. ECKBERG & MARCY R. PODKOPACZ, *FOURTH JUD. DIST. OF THE STATE OF MINN., FAMILY COURT FAIRNESS STUDY 3* (2004) (reporting that “[r]eceiving an explanation from the bench made the biggest difference” in satisfaction for trial litigants and was strongest for those who did not get what they asked for).

192. *Id.* (reporting that trial losers were more likely to say they would comply when they reported both fair treatment and having received a full explanation).

193. *Id.* at 38 (reporting that fairness and satisfaction ratings were related to whether litigants received an explanation and had an opportunity to ask questions).

194. *Id.* at 3 (describing the “explain” condition as including a question-and-answer period).

195. *See* Casey, Burke, & Leben, *supra* note 63, at 95-96 (recommending feedback and accountability loops and checklists as tools that catch errors before they harden into routine practice).

specifically at judges emphasizes that fatigue, distraction, and multitasking diminish performance, and it recommends routines that include short breaks, decision aids like checklists, and feedback loops.¹⁹⁶ Those routines complement procedural-justice values: a judge who is more mindful and focused is better able to listen carefully and to show that listening by briefly repeating key themes from the parties' testimony or argument.¹⁹⁷

Fair procedures matter most for legitimacy, but people also expect courts to establish the truth and reach substantively just outcomes, and those outcome-related goals also move legitimacy perceptions—just less than procedural justice does.¹⁹⁸ Trial-court practices that strengthen input (voice and listening) and discipline reasoning (structured reflection and checklists) do not guarantee correct outcomes, but they should improve the court's odds of getting it right—especially in the kinds of high-stakes hearings where parties are frightened, angry, or confused.¹⁹⁹

Training and implementation evidence supports cautious optimism that practices like these can improve perceptions of fair treatment. In policing, a large-scale procedural-justice training program reduced complaints and use of force over time, and the training syllabus included detailed templates for respectful interactions and de-escalation—evidence that behaviorally specified training can change real-world conduct in high-pressure, public-facing work.²⁰⁰ In Milwaukee, the courtroom-

196. *Id.* at 76-77, 93, 95-96 (identifying fatigue, mood, and multitasking as performance threats and recommending breaks, checklists, and feedback).

197. *Id.* at 77 (explaining that repeating key themes from the parties both shows listening skills and helps identify how a case differs from similar ones).

198. See Tyler & Sevier, *supra* note 12, at 1097-98 (recognizing that in addition to procedural-justice factors, whether judges punish justly and seek truth also matters to court legitimacy).

199. SHARYN ROACH ANLEU & KATHY MACK, JUDGING AND EMOTION: A SOCIO-LEGAL ANALYSIS 14-16 (2021) (describing courts as emotionally charged environments and noting the importance of practices that regulate emotion and signal fairness); see also Terry A. Maroney, *The Emotionally Intelligent Judge*, 49 CT. REV. 100, 112-113 (2013) (explaining the ways emotion is encountered in court proceedings and providing advice for trial judges handling it).

200. George Wood, Tom R. Tyler, & Andrew V. Papachristos, *Procedural Justice Training Reduces Police Use of Force and Complaints Against Officers*, 117 PROCS. NAT'L ACAD. SCIS. 9815, 9816 (2020) (reporting reductions in use of force and complaints and describing training templates for respectful, procedurally just interactions; complaints fell by ten percent and use-of-force fell by six percent).

communication intervention likewise produced measurable improvements in observed judicial behaviors (respect, fairness, attentiveness, clarity, explaining call order, assuring evidence will be considered), even though defendant perceptions and compliance did not show clear improvement—likely in part because baseline ratings were already high and the sample was limited.²⁰¹ Even if those behavioral changes by judges did not produce statistically significant results in that study, though, the study demonstrates that behavior change is measurable. Those changes, perhaps combined with others, may be a necessary precondition to genuine improvement, while sustained perception change may require reinforcement and institutional follow-through—sometimes including improvements that lie outside the judge’s control (like waiting time and procedural clarity at the courthouse front end).²⁰²

The idea of sustained change leads directly to the “by design” point. Courts have long offered one-off training sessions on procedural justice and decision-making, but many programs lack a follow-up plan that helps judges turn training into on-the-bench habits.²⁰³ The trial-court practices in this section are therefore framed as routines: short scripts, predictable explanations, visible listening moves, comprehension checks, and a small set of decision-making guardrails that can be practiced, observed, and improved over time.²⁰⁴

These trial-court practices lead to the next design problem: reasons and explanations. When court users leave a hearing, the recurring question is not only whether they were treated respectfully, but also whether they understood why the court ruled as it did and what that ruling means for them. The next section

201. See FARLEY, JENSEN, & REMPEL, *supra* note 180, at 41-42 (reporting significant improvement in courtroom behaviors but no clear change in defendant perceptions while noting high baseline ratings and sample-size limits).

202. See *id.* at 14-15, 66 (noting “front-end” predictability practices as intervention targets); KUBITS ET AL., *supra* note 178, at 3-4, 25 (showing that perceived long waits reduced ratings of fairness and that procedural confusion alters participant experiences).

203. See *Procedural Fairness in a Pandemic*, *supra* note 154, at 697 (noting that many procedural-justice trainings lack follow-up plans to turn awareness into habits).

204. See *Public Satisfaction*, *supra* note 23, at 18 (recommending concrete courtroom practices including explaining ground rules and thanking people for patience when court hearings are delayed).

therefore turns to reasons and explanations—how courts can make decision-making methods visible and intelligible to the people who must accept and follow that decision.²⁰⁵

C. Reasons and Explanations as the Bridge Between Decision Quality and Legitimacy

Reason-giving is where decision quality and legitimacy meet. Done well, explanation is one of the few judicial practices that can deliver several procedural-justice values at once—by showing voice (I heard you), neutrality and transparency (I used explainable standards), and trustworthiness (I acted in good faith rather than hidden motives).²⁰⁶

In some ways, this matters more today because the audience for opinions has changed. Appellate opinions were once written largely for lawyers, but web access means litigants and the public are now reading them in large numbers.²⁰⁷ Their sense of procedural justice increases if they can understand the explanation and feel that the court listened to both sides; it decreases if they cannot understand the decision or conclude the court did not take both sides seriously.²⁰⁸ The same point applies in trial courts and administrative settings: the process is not fair when litigants cannot understand it.²⁰⁹ And more often today, parties in the trial court are self-represented, so there is no lawyer to explain what the court said.

A good explanation also has an internal payoff. Writing for lay readers can expose flaws in legal reasoning that remain hidden

205. See generally *Getting It Right Isn't Enough*, *supra* note 168, at 33-34 (suggesting “letter to the loser” approaches and improving perceptions of fairness by making sure participants see evidence in court decisions that the court heard them).

206. See Tyler & Sevier, *supra* note 12, at 1106-07 (defining procedural justice with the elements of voice, neutrality, respect, and trustworthiness; explaining that neutrality is supported by transparency and rule-based explanations and that trustworthiness is communicated when authorities listen and explain their actions in ways showing good faith and concern). *Id.* at 1125 n.80 (noting that careful statistical analysis shows that the “four elements of procedural justice . . . are highly interrelated”).

207. *Getting It Right Isn't Enough*, *supra* note 168, at 32 (noting web access has changed the audience for court opinions, with nonlawyer readership implied by large visitor counts).

208. *Id.*

209. McCormack, *supra* note 174, at 44 (“It is not fair when litigants cannot understand it.”).

when the underlying concepts are described only in the code laypeople call legalese.²¹⁰ That is a decision-discipline point as much as a legitimacy point: an opinion that must explain itself in plain terms forces the writing judge to test whether the result makes sense at an elemental level.²¹¹ In the language of decision hygiene, it is a structured second look that can reduce the risk of something that sounds right dressed in legalese but actually skips some logical steps, isn't really on point, or is otherwise flawed in its reasoning.²¹²

Judicial-writing authorities frame the same idea as a core purpose of an opinion. A judicial opinion should help the writing judge think through the problem, justify the outcome to the parties—especially the losing litigant—and, in appellate courts, provide a legal rule citable as precedent.²¹³ The opinion-writing challenge is to justify the determination sensibly and clearly without slighting the merits on the losing side. This “losing party” focus is not merely a tone preference; it is a procedural-justice mechanism. It signals that the court took the party's arguments seriously even while rejecting them.²¹⁴

A practical model for this is the “opinion as a letter to the loser.” In one influential articulation, the court can show respect by speaking directly to the losing party, acknowledging complexity, and confronting reasons on both sides so that the decision reads as a fair resolution of competing arguments rather than a power move.²¹⁵ Consistent with that approach, therapeutic-jurisprudence writers emphasize that even when a court must rule against a party, it can still communicate that the party's arguments were fairly and fully considered, and that one concrete way to

210. *Getting It Right Isn't Enough*, *supra* note 168, at 32 (suggesting that a litigant who cannot understand an opinion cannot form a belief that the case was decided on neutral principles).

211. *Id.* at 33 (explaining that writing for lay readers often exposes flaws in reasoning).

212. See KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 298-99 (describing “decision hygiene” as preventive procedures that reduce error without requiring diagnosis of specific biases).

213. See GARNER, *supra* note 168, at § 26.1(a).

214. See *Getting It Right Isn't Enough*, *supra* note 168, at 33-34 (contrasting short affirmances that offer “no sense of validation” with short opinions that show claims were fairly heard).

215. Amy D. Ronner, *Therapeutic Jurisprudence on Appeal*, 37 CT. REV. 64, 65-66 (2000); see also *Getting It Right Isn't Enough*, *supra* note 168, at 26, 32-33.

show listening is to note at least the party's main points somewhere in the opinion.²¹⁶

Plain language should be understood as reader-based design. The aim is not merely stylistic; it is whether the text works for readers—whether they can find what they need, understand it, and use it.²¹⁷ That reader-based approach naturally supports testing. Where formal testing is possible, it has significant value, and even small readability improvements can produce large gains when documents have large audiences.²¹⁸ Courts rarely have the resources to run usability tests on every opinion, but the mindset transfers: write with the reader's tasks in mind, and use informal testing (for example, reviewing drafts with a law clerk, assistant, or colleague who takes the role of a nonlawyer reader).²¹⁹

Plain language also advances procedural fairness in a direct way. Procedural fairness includes a trust component: the decision-maker listens to litigants' concerns and explains decisions in language that is understandable.²²⁰ Former Michigan Chief Justice Bridget Mary McCormack's point is blunt and useful for judges: "It is not fair when litigants cannot understand it."²²¹ And her practical bridge is the one courts often miss: using plain language in interactions with court users is at least as important as translating forms and templates.²²² Clear explanations make fairness real.

Good explanation also benefits from predictable structure. The best writing guidance does not require judges to become writing instructors, but it does offer a few repeatable moves: lead with an umbrella paragraph that previews the roadmap, answer

216. *See id.* at 33-34 (citing Amy Ronner and Bruce Winick).

217. JOSEPH KIMBLE, *WRITING FOR DOLLARS, WRITING TO PLEASE: THE CASE FOR PLAIN LANGUAGE IN BUSINESS, GOVERNMENT, AND LAW* 17 (2d ed. 2023) (arguing that plain-language guidance aims at whether readers can "find, understand, and use" the text and that "text-based" and "reader-based" approaches are inseparable).

218. *Id.* at 18 (stating that the literature is "unequivocal" on the value of testing and that even small improvements can produce large savings where documents have large audiences).

219. *Id.* at 19 (suggesting informal testing by showing the proposed document to a client or potential reader).

220. McCormack, *supra* note 174, at 45 (defining "trust" as including listening and explaining decisions in understandable language).

221. *Id.* at 44.

222. *Id.* at 45.

the questions a skeptical reader will naturally ask, and acknowledge viable counterarguments rather than pretending they do not exist.²²³ Relatedly, tone is not an afterthought. A judicial writer should be able to fairly answer these questions: “Does the losing party get a fair shake?” and “Have I resisted the temptation to slant the facts?”²²⁴

Finally, explanations link back to intellectual humility in a concrete way. Intellectual humility becomes visible when, rather than just announcing and supporting one view, the court acknowledges complexity and confronts competing reasons.²²⁵ But that visibility depends in part on the institution’s willingness to permit disagreement. In a group setting, psychology professor Charlan Nemeth’s synthesis of the dissent literature stresses that consensus narrows and closes thinking while dissent broadens it; even a single dissenter can break the power of unanimity and increase independent thinking.²²⁶ Similar methods are available even to a solo trial judge: a premortem and other second-look practices can test an emerging conclusion before the decision is rendered.²²⁷ The next subsection turns to the collective side of design—how courts can protect candid disagreement while improving the clarity and legitimacy of the reasons they ultimately give—while also noting how individual judges

223. ROSS GUBERMAN, *POINT MADE: HOW TO WRITE LIKE THE NATION’S TOP ADVOCATES* 107-110 (2d ed. 2014) (recommending the use of umbrella paragraphs before headings and subheadings to orient the reader and question-driven or numbered-reason organization to guide analysis).

224. *Id.* at 23, 30 (2015) (advising “[e]dit for tone” and asking whether the losing party gets a “fair shake” and whether the writer has resisted slanting).

225. *See id.* at 23 (explaining that decisions can be well-received when they acknowledge complexity and confront reasons on both sides).

226. CHARLAN NEMETH, *IN DEFENSE OF TROUBLE MAKERS: THE POWER OF DISSENT IN LIFE AND BUSINESS* 2, 39 (2018) (explaining that a single dissenting voice can increase independent thinking of a group).

227. Burns & Winthrop, *supra* note 151 (defining a premortem and explaining that it reduces overconfidence and social-desirability pressure by making it psychologically safer to surface doubts); KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 43, at 298-99 (explaining decision hygiene procedures that reduce error); Krumrei-Mancuso et al., *supra* note 5, at 22 (listing premortems and second-read policies as institutional scaffolds enabling humility at scale).

(whether at the trial or appellate level) can incorporate similar processes within chambers.²²⁸

D. Collective Practices: Appellate Courts, Chambers, and Court Administration

Judicial decision-makers do not always work alone. Much of judging is collegial—appellate judges sit in panels and court administration often involves committees and group work. Federal trial judges are supported in chambers by law clerks (law-school graduates) who perform legal research and writing, including drafting orders and opinions and serving as a confidential sounding board for their judges.²²⁹ In state trial courts, chambers staffing varies by jurisdiction, but many judges have confidential administrative chambers support (often a nonlawyer judicial assistant), and some jurisdictions provide support through either law student clerks or law graduates, sometimes organized as shared or team-based resources, to help draft and complete case-management work.²³⁰ This provides

228. Krumrei-Mancuso et al., *supra* note 5, at 23 (describing intellectual humility); *Why (and When) Judges Dissent*, *supra* note 150, at 102 (defining “dissent aversion” and discussing conditions that limit dissent in appellate courts).

229. *Legal Careers*, U.S. CTS., [https://perma.cc/ZA87-FR2G] (last visited Mar. 29, 2026) (describing federal judiciary law clerk positions).

230. See generally RON MALEGA & THOMAS H. COHEN, U.S. DEP’T OF JUST., STATE COURT ORGANIZATION, 2011 8 tbl. 8 (2013), [https://perma.cc/438P-8HCV] (stating trial-court “research attorney salaries” tracked as an expenditure item with mixed state/county funding); *State Court Organization*, tbl. 14, [https://perma.cc/2329-DWFB] (last visited Feb. 11, 2026) (including “q5 Research Attorney/Law Clerk” within trial-court salary/expenditure lines); see also, e.g., ALASKA CT. SYS., CLASS SPECIFICATION: CHAMBERS JUDICIAL ASSISTANT 1 1-2, [https://perma.cc/BPF8-CGZA] (last visited Feb. 11, 2026) (defining chambers judicial assistant as confidential administrative and technical support for one or more trial judges and listing drafting and calendar duties); CIR. CT. FOR MONTGOMERY CNTY., MD., PERSONNEL CLASS SPECIFICATION: JUDICIAL ASSISTANT 1-2, [https://perma.cc/ZB2B-LHEV] (last visited Feb. 11, 2026) (noting confidentiality and listing case-calendaring and “in-chambers conferences”); N.D. CT. SYS., CLASSIFICATION DESCRIPTION: STAFF ATTORNEY-DISTRICT COURT, [https://perma.cc/88L9-Z7YA] (last visited Feb. 11, 2026) (describing J.D.-required chambers attorney drafting orders for multiple judges and supporting “in-chamber conferences”); CUYAHOGA CNTY. COMMON PLEAS CT. GEN. DIV., JOB DESCRIPTION 1 (2024), [https://perma.cc/YR6U-GUJ2] (describing J.D.-required staff attorney assisting “one or two Judges” with docket management, drafting, and case-management conferences); MEGAN LAVOIE, OFF. CT. ADMIN, JOB POSTING (2025), [https://perma.cc/N5LS-KJ62] (describing J.D.-required staff attorney preparing draft rulings/opinions under a confidentiality policy); N.J. Cts., *Law Clerk 2026-2027* (June 4, 2025),

additional opportunities to infuse intellectual-humility habits into judicial routine. Intellectual humility is both an individual attribute and a collective property that can be supported or undermined by institutional design.²³¹ In a group setting, professional norms and incentives can push participants in a group toward quick agreement, even when candor and reflective thinking would improve the quality of the decision.²³²

The basic cognitive risk here is that consensus can narrow thinking. Professor Nemeth's synthesis of the dissent literature emphasizes that consensus often changes how people think in closed and narrow ways, while dissent broadens thinking and improves the quality of group decisions.²³³ Even one dissenter—think Henry Fonda as Juror No. 8 in *Twelve Angry Men*²³⁴—can break the power of unanimity and increase independent thinking. The point is not celebration of the Henry Fonda characters in the world; it's what happens to the thinking process when at least one person dissents. Disagreement forces people to search for reasons, consider alternatives, and revisit assumptions that might otherwise go untested.

Appellate courts face predictable structural pressures that can suppress this kind of productive disagreement. Professor Lee Epstein, Judge Richard Posner, and economist William Landes described “dissent aversion,” a tendency for judges not to dissent even though they disagreed with a majority of an appellate court panel.²³⁵ They explained that dissents take time and effort, and they come with another potential cost—lost collegiality.²³⁶ If the goal is to get the best possible exploration of issues and, based on

[<https://perma.cc/XD7V-KUF3>] (listing law school graduation requirement; trial court clerk duties include summarizing information for judges and emergent-motions support).

231. Krumrei-Mancuso et al., *supra* note 5, at 15, 23 (arguing that collective intellectual humility depends on norms, roles, procedures, and scaffolds, not merely the average humility of members).

232. Porter et al., *supra* note 112, at 528 (noting how group conformity and status motives can work against intellectual humility).

233. NEMETH, *supra* note 226, at 80 (stating that consensus changes thinking in narrow and closed ways while dissent broadens thinking and improves decision quality).

234. 12 ANGRY MEN, Amazon Prime (Orion-Nova Prods. 1957).

235. *Why (and When) Judges Dissent*, *supra* note 150, at 101-03 (defining “dissent aversion” and reporting that dissent frequency is negatively related to caseload and positively related to ideological diversity).

236. *Id.* at 104-05.

that, the best possible court decision, the institutional design condition seems straightforward: find ways to encourage and allow constructive disagreement between judges who work together.

Decision-hygiene research provides another lens for better decision-making.²³⁷ Noise (unwanted variability in judgments) is often invisible unless an institution measures it, carefully studies it, and manages it.²³⁸ Kahneman, Sibony, and Sunstein suggest that careful management through decision-hygiene procedures is preventive: it reduces error without requiring the decision-maker to diagnose which bias might be operating in a particular case.²³⁹ Noise-audit responses can take different forms depending on the setting—from aggregating independent judgments before discussion in some professional environments to more formal constraint-based approaches in courts, such as sentencing guidelines to reduce disparities.²⁴⁰

Incorporating decision hygiene into collective court processes could have similar advantages. It reframes “second looks” as professional method rather than personal weakness. It also supplies a practical vocabulary for court leaders: we are not accusing anyone of bias; we are improving process to reduce inconsistency and error.

While decision hygiene supplies the process tools, scaffolding supplies the institutional supports that make these tools routine rather than aspirational. There are several low-cost scaffolds that fit well in court settings. A premortem asks the team to assume that a decision or plan has failed and work backward to identify why.²⁴¹ It is designed to counter overconfidence and social-desirability pressure by making it psychologically safer to voice doubts, even ones that seem to come from unlikely

237. Krumrei-Mancuso et al., *supra* note 5, at 19-22 (describing implementation mechanisms such as norms and incentives that determine whether disagreement improves decisions or is suppressed); *see also*, Tenelle Porter et al., *Intellectual Humility Predicts Master Behaviors When Learning*, LEARNING & INDIVIDUAL DIFFERENCES, May 2020, at 1, 2.

238. KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 15.

239. *Id.* at 298-99.

240. *See id.* at 181.

241. Gary Klein, *Performing a Project Premortem*, HARV. BUS. REV., Sept, 2007, at 18, 18-19.

possibilities.²⁴² Think of the famous *Chicago Daily Tribune* headline, “DEWEY DEFEATS TRUMAN.” Their subhead had it right: “REPUBLICAN TICKET AHEAD OF 1944 VOTE.”²⁴³ The subhead said nothing about the final result. But the top headline was big and bold—and wrong. A premortem—if this turns out to be wrong, what would be the reason?—might have saved the paper national embarrassment.²⁴⁴ But it would have deprived history of one of the all-time great presidential photos: President Truman beaming and holding that paper up for all to see.²⁴⁵

Groups can use premortems and second-read policies to enable intellectual humility at scale. These procedures make doubt and error-surfacing routine rather than exceptional.²⁴⁶ In the appellate context, second-look practices can take familiar forms with an emphasis on balanced presentation: preconference case summaries designed to highlight the strongest arguments for the losing party, structured checklists for recurring decision points, and explicit prompts that require the panel to articulate what would make its preferred outcome wrong. Even if the result stays the same, the discussion may lead to a better opinion.

242. See generally *id.* (introducing the premortem as a planning exercise in which a team assumes the project has failed and generates plausible reasons for the failure); BETH VEINOTT, GARY A. KLEIN, & STERLING WIGGINS, EVALUATING THE EFFECTIVENESS OF THE PREMORTEM TECHNIQUE ON PLAN CONFIDENCE 1-3 (2010) (defining a premortem as proceeding “[on] the premise that a program or plan or project has failed[,]” outlining steps for generating reasons for failure, and explaining why the method may reduce overconfidence through “prospective hindsight”); Guillermo M. Wippold et al., *Proactive Planning for Contextual Fit: The Role of the Implementation Premortem*, IMPLEMENTATION SCI. COMM’NS, Aug. 7, 2025, at 1, 3 (describing an implementation premortem protocol that asks participants to assume program failure and brainstorm reasons for failure and strategies to address them). For a plain-language discussion, see generally Burns & Winthrop, *supra* note 151 (explaining how premortems surface risks and reduce overconfidence and social-desirability pressures).

243. Eric Allen, *Stop the Presses!*, LIBRR. OF CONG. BLOG (Nov. 6, 2012), [https://perma.cc/6GDH-KTGV].

244. See Tim Jones, *Dewey defeats Truman: The most famous wrong call in electoral history*, CHI. TRIB. (Oct. 31, 2020, at 6:57 PM), [https://perma.cc/XBL2-QMHY] (“Arguably the most famous headline in the newspapers 150-year history, DEWEY DEFEATS TRUMAN is every publisher’s nightmare on every election night.”).

245. See Ben Cosgrove, *Behind the Picture: ‘Dewey Defeats Truman’*, LIFE (“The reason for the picture’s immortality? . . . the picture endures because of the look of unabashed, in-your-face delight in Truman’s eyes.”), [https://perma.cc/82MT-Z52V] (last visited Feb. 23, 2026).

246. See Krumrei-Mancuso et al., *supra* note 5, at 22.

This design can be built into both preconference review and the mechanics of deliberation. A simple rule—“someone must summarize the best version of the other side’s argument”—is a humility practice that should improve both accuracy and perceived legitimacy. It also reduces the risk that disagreement becomes personal conflict rather than information.²⁴⁷ After all, if someone needs to state the strongest contrary view, the mere act of doing so becomes part of the process, not a challenge to those who already have a firmly held view on the other side.

Building in the space to disagree has another advantage. Research on listening shows that when people feel heard, they become less defensive and more reflective, increasing self-insight and reducing extreme attitudes.²⁴⁸ That dynamic matters in conference and chambers. Dissent that is heard is more likely to improve the decision; dissent that is treated as disloyalty is more likely to disappear.

Many of these ideas can be applied not only to multi-judge panels but also to solo trial judges working in chambers. A trial judge can harness the value of dissent by creating it internally—through structured self-critique before signing an order. Self-distancing research suggests that people reason more wisely about others’ problems than their own, but instructing someone to adopt a self-distanced perspective eliminates that self-other asymmetry.²⁴⁹ One way to do this is to change vantage point before finalizing a ruling. For example: “If I were an appellate judge reviewing this order, what is the strongest concern I would have?” Or: “If I were advising another judge on this case, what would I tell them to double-check before signing?” A premortem can serve a related function: “Assume this decision is wrong—what would the most likely reason be?”²⁵⁰ In addition, trial judges

247. See Porter & Schumann, *supra* note 115, at 159 (linking intellectual humility with openness and constructive engagement with opposing views).

248. See Itzhakov et al., *supra* note 141, at 213, 237 (theorizing and testing a pathway from high-quality listening to positivity resonance, self-insight, and reduced attitude extremity).

249. Grossman & Kross, *supra* note 138, at 1571-72, 1577 (reporting that people reason more wisely about others’ conflicts than their own and that instructing a self-distanced perspective eliminates the self-other asymmetry).

250. See VEINOTT, KLEIN, & STERLING, *supra* note 242, at 3 (explaining how premortems surface risks).

who have law clerks or confidential employees can develop in-chambers systems with the same tools an appellate judge might use to let structured discussion in a group setting raise potential issues about a tentative outcome on an issue or case.

Collective design also helps to overcome the limits that come by focusing only on improving performance on one or two psychological traits. For example, actively open-minded thinking closely overlaps with intellectual humility in the practices it prescribes—seeking counterevidence, delaying premature closure, and revisiting when reasons warrant. But measurement studies report that actively open-minded thinking scales do not reliably predict myside bias, and some findings suggest counterintuitive interactions with ideology in polarized settings.²⁵¹ That suggests the importance of design over focus on a trait: design procedures can force the search for counterarguments and institutionalize second looks in a more reliable way than exhortations to be open-minded can.

Finally, court administration can help make collective practices durable. Courts can institutionalize humility through feedback loops, observation, and accountability mechanisms that make it easier to spot drift or inconsistency and reinforce good habits. Judicial decision-making research emphasizes that fatigue, distraction, and multitasking degrade performance and that decision aids—like checklists and feedback—can improve consistency and judgment.²⁵² These are administrative choices as much as individual ones. Courthouse design now unwittingly encourages multitasking and distraction by the computers (often

251. Stanovich & Toplak, *supra* note 132, at 160-61 (reviewing multiple studies and concluding that, despite predicting many other biases, actively open-minded thinking questionnaire scales show negligible or non-significant relationships with direct measures of myside thinking); Kahan & Corbin, *supra* note 136, at 1-2 (reporting that, based on a national survey, polarization on climate-change beliefs increases as actively open-minded thinking scores increase because higher-AOT liberals move toward greater acceptance while higher-AOT conservatives do not).

252. Casey, Burke, & Leben, *supra* note 63, at 76-77, 95-96; *see also* Haynes et al., *A Surgical Safety Checklist to Reduce Morbidity and Mortality in a Global Population*, 360 NEJM 491, 492-98 (2009) (reporting that implementation of a nineteen-item surgical safety checklist designed to improve team communication and consistency of care was associated with lower death and complication rates across eight hospitals worldwide).

more than one) on the courtroom bench.²⁵³ Beyond physical design, court administration, led by judges, can set expectations for practices that combine intellectual humility, decision science, and procedural-justice lessons. A court can, by design, treat candor, second looks, and structured disagreement as institutional expectations—and support that by routines that make those practices easier to do when calendars are crowded and stakes are high.

E. Why Design Matters for Institutional Change

The justice system's institutional values include independence, impartiality, and equal treatment,²⁵⁴ and it's easy to agree in principle with procedural-justice principles. The implementation challenge is converting those values into reliable habits—especially under time pressure, fatigue, and lots of cases on the daily court calendar. That is why design matters: one-off training can raise awareness, but without reinforcement and follow-through, new practices often do not persist in day-to-day judicial work. And, as with decision hygiene generally, the benefits of preventive routines are often invisible—difficult to trace to any single avoided error—so they are harder to sustain unless the institution makes them easy and expected.

Design also clarifies what courts are trying to change. Abstract commitments (“be fair,” “be impartial,” “be open-minded”) are not readily observable. But many procedural-justice and decision-quality practices can be operationalized as concrete behaviors: explaining what will happen next, making visible eye contact, using plain English to explain procedure, asking whether a party has anything to add before a decision, and checking understanding of next steps. In addition, when specific behaviors are involved, measurement is possible. When courts define

253. See Kevin Burke & Steve Leben, *Procedural Fairness*, in ENHANCING JUSTICE: REDUCING BIAS 229, 233-34 (Sarah E. Redfield ed., 2017) (discussing problems of perception and divided attention due to computers on courtroom benches).

254. See MODEL CODE JUD. CONDUCT pmbl. ¶ 1 (Am. Bar Ass'n. 2011) (“An independent, fair and impartial judiciary is indispensable to our system of justice.”); *id.* Canon 1 (requiring judges to uphold and promote the “independence, integrity, and impartiality of the judiciary”); *id.* r. 2.3(A)–(B) (requiring judges to perform judicial duties without bias or prejudice and prohibiting manifestations of bias or prejudice).

practices in observable terms, they can gather feedback—through court-user surveys, periodic observation, peer coaching, or video review—and use that feedback to identify drift and reinforce improvement and good performance.

The Milwaukee courtroom-communication experiment illustrates both the promise and the limits of a design approach. The intervention combined a one-day training with individualized action plans, and structured courtroom observations detected significant improvements in many targeted communication behaviors.²⁵⁵ But that study did not find evidence that defendants formed more positive perceptions due to those observed changes, and official compliance and recidivism measures did not show significant improvements within the study's time frame.²⁵⁶

Even so, there is strong causal evidence in other court studies of the positive impacts of procedural justice on court legitimacy.²⁵⁷ There has also been extensive work—outside courts but directly transferable as an institutional-change model—that training built around procedural justice can produce measurable reductions in harmful conduct. In Chicago, for example, a procedural-justice training program for thousands of officers reduced use of force over two years, and the program emphasized respect, neutrality, transparency, and opportunities for people to explain their side.²⁵⁸ The court-facing lesson is that

255. FARLEY, JENSEN, & REMPEL, *supra* note 180, at iii-iv, 24-26 (describing court program that included one-day training with individualized judicial action plans and reporting that resulted in statistically significant improvement in targeted communication behaviors like more frequent use of explanatory opening practices, increased use of plain English, more frequent invitations for defendants and defense counsel to speak before decisions in criminal cases, and increased observable indicators of attention to comprehension).

256. *Id.* at iv-v, 27-28 (finding no evidence that defendants formed more positive perception due to the observed changes in communication and finding no statistically significant differences in nine-month compliance or four-month recidivism time frames within the study windows).

257. *Getting It Right Isn't Enough*, *supra* note 168, at 22-23 (summarizing research showing that procedural fairness drives views of the justice system and compliance with court orders); Kevin Burke & Steve Leben, *The Evolution of the Trial Judge from Counting Case Dispositions to Commitment to Fairness*, 18 WIDENER L. J. 397, 404 n.28 (2009) [hereinafter *Evolution of Trial Judges*] (citing research showing that procedural justice is more important in perception of court experience than winning or losing).

258. Wood, Tyler, & Papachristos, *supra* note 200, at 9815, 9817 (describing a Chicago Police Department rollout of a one-day procedural-justice training and reporting reductions in use-of-force reports over a two-year after-training period).

large public-facing institutions can change observable behavior when the training is behaviorally specified and supported by institutional follow-through rather than moral exhortation alone.²⁵⁹ Tying together related techniques from separately studied fields—procedural justice, decision science, and intellectual humility—should increase the chances that courts will function better and that participants will have a better experience.

Finally, design matters because court practices do more than move cases along; they also communicate what judging is *for*. As law professor Cass Sunstein argues, law works not only by changing incentives but also by sending messages that shape social norms—but the value of those messages isn’t automatic.²⁶⁰ Put in court-system terms, Sunstein’s point is that an official statement should be judged not simply by whether it affirms the right values, but by whether it actually helps shape public understanding, conduct, and confidence in the institution.²⁶¹ The same point applies to courts as institutions. A court that builds routines of explanation, transparency, and candid self-checking is not merely improving internal decision-making; it is making a public statement about what counts as professional judging. And Sunstein’s caution matters too: expressive messages can fail or backfire when the audience hears something different from what the institution intended—especially when communication is opaque, jargon-heavy, or misleading.²⁶² That is another reason to treat procedural fairness as design: the goal is not only to do the right thing, but to make it legible—so court users can see what the court did and why, while also concluding the court did its best to be fair.

259. *Id.* at 9816, 9819-20 (explaining that police retraining is often invoked but rarely evaluated and arguing that this large-scale procedural-justice training provides evidence that realistic training can change on-the-job behavior and support broader institutional culture change).

260. Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021, 2025-26 (1996).

261. *Id.* at 2021-22, 2025 (framing the expressive account as one that must be judged by the “ultimate effects” on institutions and that empirical questions about effects must be answered).

262. *Id.* at 2022, 2050-51 (warning that government efforts can be counterproductive when the audience hears a different message and that the meanings of actions are not fully within an agent’s control).

For courts that want change to stick, then, the implementation playbook is practical: choose routines that are easy to repeat; define them behaviorally; measure them; and reinforce them. The Milwaukee judges' own recommendations fit that template—provide research in advance, use evidence to persuade skeptical judges, include booster sessions, and involve the whole courtroom team so that communication norms become shared practice rather than individual style.²⁶³ Courts can make the same move with decision hygiene tools: simple checklists and real-time observation can prevent predictable mistakes and help ensure that “fair process and just outcomes” are pursued together.²⁶⁴ Part VI turns from design principles to adoption and evaluation—how courts can pilot these practices, measure whether they are working, and revise them with intellectual humility when they are not.

VI. IMPLEMENTING INTELLECTUAL HUMILITY BY DESIGN

Earlier sections have explained how intellectual humility connects procedural justice and decision science with court practice and court administration. This section seeks to embed those practices into typical judicial practices. It describes some bench tools, templates, and workflow prompts that, based on the research summarized here, should be helpful. It also addresses the limits of these tools and suggests research that could test and improve these materials. Since these tools have not yet been tested in courts, they are intended as starting points based on current evidence, not fully vetted final products.

263. FARLEY, JENSEN, & REMPEL, *supra* note 180, at 41 (recommending various implementation supports, including advance reading, emphasizing evidence to persuade skeptical judges, adding a booster session, and including the entire courtroom team so that procedural justice becomes a shared practice).

264. Casey, Burke, & Leben, *supra* note 63, at 76-78, 95-96 (citing effects of fatigue, multitasking, and distraction on performance; recommending decision aids like checklists and feedback); *see also* Haynes et al., *supra* note 252, at 495-98 (noting success of checklist usage in improving surgical outcomes).

A. An Appendix of Sample Tools and Pathways to Adoption

If intellectual humility is to become “by design” rather than episodic, it must be made usable in ordinary court conditions—crowded calendars, high emotional stakes, imperfect information, and constant time pressure. Tools should therefore meet three design constraints: they must be (1) short enough to be used in real time, (2) timed to real decision points in the work (before court, at the outset of a hearing, during argument or testimony, when announcing a ruling, and after the matter concludes), and (3) behaviorally specific enough that the judge and court staff can see what adopting the tool means.²⁶⁵

The Appendix begins with a bench card that operationalizes these principles: the Intellectual Humility by Design Bench Card.²⁶⁶ It has four parts: (1) Recognize fallibility;²⁶⁷ (2) Reflect

265. See Casey, Burke, & Leben, *supra* note 63, at 76-77, 95-96 (discussing time pressure, distraction, and the utility of decision aids like checklists and bench cards).

266. See *infra* p. 213.

267. See *infra* p. 213. The “Recognize fallibility” prompts (what would change my mind; consider the opposite; name uncertainty; outside view) are grounded in intellectual humility research defining the construct as recognizing fallibility and limits of knowledge and linking it to reduced bias, better engagement with counterevidence, and related epistemic behaviors. See Porter et al., *supra* note 112, at 524-26 (synthesizing definitions and measurement approaches and describing intellectual humility as a metacognitive core that recognizes limits and fallibility); Leary et al., *supra* note 122, at 794-97 (defining intellectual humility as recognizing one’s beliefs might be wrong and linking it to argument evaluation and reduced defensiveness). The “premortem” and “if criticized or reversed, why?” prompts are supported as structured “second-look” techniques designed to counter overconfidence and social desirability pressure. See Kirmrei-Mancuso et al., *supra* note 5, at 22; VEINOTT, KLEIN, & STERLING, *supra* note 242, at 7; Guillermo M. Wippold et al., *supra* note 242, at 4; KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 370 (describing “decision hygiene” as preventive procedures that reduce error without diagnosing a particular bias).

before ruling;²⁶⁸ (3) Respond with reasons;²⁶⁹ and Reassess after decision.²⁷⁰ It is drafted as a companion to the widely used

268. See *infra* p. 213. The “Reflect before ruling” prompts operationalize procedural justice—voice, neutrality, respectful treatment, and trustworthiness—as mechanisms for legitimacy and decision acceptance. See *Procedural Justice and the Courts*, *supra* note 17, at 26-28 (describing the four-factor model and its relationship to acceptance and legitimacy); *Public Satisfaction*, *supra* note 22, at 5-6 (same).

For the “voice first / here is what I hear—did I get that right?” and “paraphrase and invite a one-sentence correction” techniques: these reflect procedural-justice “voice” as a perception that the authority sincerely considered what the party said and with empirical work on the effects of listening quality on defensiveness and reflection. *Procedural Justice and the Courts*, *supra* note 17, at 28, 30 (explaining voice as meaningful opportunity to be heard and considered); see also Itzhakov et al., *supra* note 141, at 215, 219, 220 (finding that high-quality listening can reduce attitude extremity through positivity resonance and self-insight). For the “pause / take under advisement if tempers rise” and “name the process move” prompts: these reflect research showing the role of emotion and attention in judgment and the need for deliberate slowing under pressure. See Casey, Burke, & Leben, *supra* note 63, at 93-96 (describing fatigue/distraction and recommending breaks, checklists, and feedback routines); ANLEU & MACK, *supra* note 199, at 12-13 (describing courts as emotionally charged environments and the importance of practices that support careful, respectful decision-making); Terry A. Maroney, *Emotional Regulation and Judicial Behavior*, 99 CAL. L. REV. 1485, 1486, 1510-11, 1534 (2011) (explaining that emotion regulation and attention control draw on limited cognitive resources; that suppression increases cognitive load and impairs memory; and that judges sometimes should take a break or step away to calm down and think before responding).

269. See *infra* p. 213-15. The “Respond with reasons” prompts are anchored in the procedural-justice literature’s emphasis on neutrality/transparency and trustworthiness cues, and in judicial-writing authority emphasizing justification to the losing party. *E.g.*, Tyler & Sevier, *supra* note 12, at 1106-07 (explaining that neutrality requires transparency and explanation, and that trustworthiness cues are communicated when authorities listen and justify actions with sensitivity). See also GARNER, *supra* note 168, at § 26.1(a) (stating that opinions should justify outcomes, especially to the losing litigant, and should not slight the losing side’s merits); McCormack, *supra* note 174, at 44-45 (arguing that courts are not fair when litigants cannot understand them and urging plain language in explanations).

The “address the strongest opposing point (as the party would state it)” and “best contrary authority respectfully” prompts are supported in the appellate procedural-justice literature: showing that key arguments were heard and taken seriously increases perceived procedural justice. *Getting It Right Isn’t Enough*, *supra* note 168, at 33-34 (explaining that procedural-justice perceptions are enhanced when readers understand the explanation and feel the court listened to both sides; discussing “letter to the loser” approaches).

Plain-language and reader-centered design support for the “roadmap early” and “plain-language layer” ideas is strengthened by empirical and practice-based writing literature. KIMBLE, *supra* note 217, at 17-19 (explaining reader-based plain language and advocating testing); see also Ash et al., *Translating Legalese: Enhancing Public Understanding of Court Opinions with Legal Summarizers*, CSLAW ‘24: PROCS. 2024 SYMP. COMP. SCI. & L., Mar. 12-13, 2024, at 136, 142-43 (reporting that simplified summaries improved understanding and perceived clarity/balance, especially for non-college readers).

270. See *infra* p. 214. The “Reassess after decision” prompts are supported by (a) decision-hygiene logic that error reduction is preventive and process-based and (b) group-process research showing disagreement broadens thinking and reduces premature consensus.

Procedural Fairness/Procedural Justice Bench Card for trial judges.²⁷¹ The procedural-justice bench card offers practical courtroom prompts tied to voice, neutrality, respect, trust, understanding, and helpfulness.²⁷² The Intellectual Humility by Design Bench Card is designed to fit inside that same practical space: it provides micro-prompts and short scripts that make

See KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 370-74 (describing preventive “decision hygiene”) (italics removed); NEMETH, *supra* note 226, at 39 (explaining that dissent broadens thinking and that a single dissenter can break unanimity pressure).

The “invite perspective; treat disagreement as data” and “assign a rotating counterargument role” prompts reflect collective intellectual humility and dissent research: groups are more accurate and less error-prone when procedures make dissent and second looks safe and expected. Kurmrei-Mancuso et al., *supra* note 5, at 18-23 (describing group-level humility and scaffolds such as roles, norms, and second reads); see also *Why (and When) Judges Dissent*, *supra* note 150, at 2-3, 6 (defining dissent aversion and describing costs that suppress dissent in collegial courts).

The “quarterly R4 retro” (reviewing reversals, unclear passages, template tweaks) is supported by some research involving judges and by research in other professions. Regular retrospective review routines—often studied as “audit and feedback,” structured debrief, or after-action reviews—tend to produce small-to-moderate improvements on average, though effects vary widely and depend on design. See N. Ivers et al., *Audit and Feedback: Effects on Professional Practice*, COCHRANE DATABASE SYST. REV., Mar. 25, 2025, at 2-3 (reporting typical modest absolute improvements (median ~2.7%; mean ~6.2%) and identifying design features associated with larger effects, including credible/benchmarked feedback, interactive delivery, and specific action planning). Meta-analyses of structured debriefs likewise report average performance improvements, especially when the review is structured and aligned to the level of performance being improved. See Scott I. Tannenbaum & Christopher P. Cerasoli, *Do Team and Individual Debriefs Enhance Performance? A Meta-Analysis*, 55 HUM. FACTORS 231, 237-41 (2013) (reporting overall positive performance effects and moderators tied to structure/alignment); Nathanael L. Keiser & Winfred Arthur, Jr., *A Meta-Analysis of the Effectiveness of the After-Action Review (or Debrief) and Factors That Influence Its Effectiveness*, 106 J. APPLIED PSYCH. 1007 (2021) (reporting a positive overall effect and emphasizing moderators, including alignment and facilitation features). In courts specifically, causal evidence suggests trial judges adjust later sentencing in the direction of appellate reversals—suggesting that outcome review can function as feedback—while also cautioning that responses may overreact to noisy signals, underscoring the value of a structured “retro” that focuses on one measured, actionable improvement rather than ad hoc reactions. See Manudeep Bhuller & Henrik Sigstad, *Feedback and Learning: The Causal Effects of Reversals on Judicial Decision-Making*, 92 REV. ECON. STUD. 2359, 2359-63 (2025). As an implementation practice for judges, see also Casey, Burke, & Leben, *supra* note 63, at 95-96 (urging feedback and accountability mechanisms and cautioning that improvement requires information about what is and isn’t working).

271. See *infra* note 272 and accompanying text.

272. The procedural-justice bench card was published in 2017 as a joint product of the American Judges Association (“AJA”), National Center for State Courts, National Judicial College, and Center for Court Innovation in the AJA’s publication, *Court Review*. Am. Judges Ass’n et al., *Procedural Fairness/Procedural Justice: A Bench Card for Trial Judges*, 53 CT. REV. 593 (2017), [<https://perma.cc/8LUQ-Z384>].

intellectual humility visible (to litigants, lawyers, staff, and the judge) and that create small “speed bumps” against rushed or automatic decision-making.

This new bench card is organized around the workflow of judging rather than around abstract categories. For trial work, the bench card targets five moments: a brief self-check before the calendar begins; an opening script that frames the process in terms parties can understand; mid-hearing prompts that encourage careful listening and accurate summarizing of each side’s best points; a ruling structure that addresses the losing party with respect and gives transparent reasons; and a short “after” note that captures what the judge almost missed or would do differently next time.²⁷³ For appellate work, it targets the same functional equivalents: preparation before argument, disciplined questioning at argument, structured conference practices, and drafting habits that keep the opinion modest, clear, and responsive to the losing party’s strongest contentions.²⁷⁴

Several features of the Intellectual Humility by Design Bench Card are intentionally mundane. Other professions have learned that decision aids—checklists and bench cards—can improve performance precisely because they are simple, repeatable, and usable when the decision-maker is tired, distracted, or under pressure.²⁷⁵ Judges may resist such tools for the same reasons physicians once resisted surgical checklists: “we are too busy,” “we already know this,” and “this will slow us down.”²⁷⁶ But the appellate perspective is instructive here. When a case returns years later because a basic step was missed, no one benefits—not the parties, not the trial court, and not the system’s legitimacy.²⁷⁷

Because the Intellectual Humility by Design Bench Card is designed as a companion to procedural fairness, it also fits comfortably within existing judicial education and court-administration structures. State-court leadership has embraced

273. *See infra* p. 215.

274. *See infra* pp. 215-16.

275. *See* Casey, Burke, & Leben, *supra* note 63, at 94-95 (describing evidence from other professions supporting checklists and decision aids).

276. *Id.* at 95 (noting objections that checklists can slow work or are unnecessary).

277. *Id.* (observing that reversals for missed steps waste time and serve no one).

procedural-justice principles for more than a decade.²⁷⁸ Many courts already teach procedural-fairness principles and encourage judges to use short scripts: introduce yourself, greet parties neutrally, acknowledge time pressures, explain procedures, and show—explicitly—that you listened.²⁷⁹ Intellectual Humility by Design does not compete with those practices; it supplies additional prompts aimed at cognitive discipline (e.g., reasons-first sequencing, “consider-the-opposite” checks, and accurate summaries of each side’s best points) that can strengthen both perceived fairness and substantive decision quality.

In addition to the bench card, four standalone tools can be adopted independently or used together. First, the Reasons and Explanations Toolkit²⁸⁰ provides plain-language prompts and a disciplined structure for explaining outcomes—especially to the losing party—while keeping the court’s reasoning transparent and limited to the decision points that control the result. Second, the

278. The Conference of State Court Administrators endorsed the American Judges Association white paper on procedural fairness in 2008, and the Conference of Chief Justices in 2013 urged state courts to incorporate procedural-justice principles into their judicial-education programming. *Getting It Right Isn’t Enough*, *supra* note 168, at 22. The AJA white paper was published in 2007 as *Public Satisfaction*, *supra* note 22.

279. *See* Am. Judges Ass’n et al., *supra* note 272 (urging introductions, neutral greetings, addressing timing concerns, and explaining procedures).

280. *See infra* pp. 217-19. The Reasons-and-Explanations Toolkit is supported by procedural-justice research showing that litigants’ perceptions of fairness and legitimacy depend heavily on whether the judge communicates neutrality, respect, and trustworthy motives—especially through clear explanations and visible listening. *E.g.*, *Procedural Justice and the Courts*, *supra* note 17, at 27-29 (synthesizing evidence that perceived procedural justice predicts legitimacy and acceptance); *Public Satisfaction*, *supra* note 22, at 6 (explaining that explanation, voice, neutrality, and respect drive public satisfaction and acceptance). The Toolkit’s emphasis on plain language and sequencing also draws on court-focused communication evidence that concrete, observable communication behaviors (including explaining what will happen and why) can be taught and implemented in real court settings. FARLEY, JENSEN, & REMPEL, *supra* note 180, at 24-26 (reporting measured changes in judge communication behaviors after training). The Toolkit’s “plain-language layer” and reasons discipline adhere to longstanding judicial-writing guidance and plain-language authority. GARNER, *supra* note 168, at § 26.1-26.5 (providing style and organization guidance aimed at clarity and reader comprehension); KIMBLE, *supra* note 217, at 6-10 (providing plain-language drafting principles and organization techniques); McCormack, *supra* note 174, at 44-45 (urging plain language as an access-to-justice imperative); *see also* MARIA MINDLIN, IS PLAIN LANGUAGE BETTER? A COMPARATIVE READABILITY STUDY OF COURT FORMS, [https://perma.cc/4C4B-WF8T] (last visited Mar. 19, 2026) (reporting improved comprehension when court forms are converted to plain language).

Trial Oral-Ruling Sequencing Options²⁸¹ tool offers formats for delivering oral rulings that improve understanding and reduce unnecessary escalation, including a two-pass option for high-stakes or emotionally-charged matters and short comprehension checks for next-step obligations. Third, the Constructive

281. See *infra* pp. 220-24. The Trial Oral-Ruling Sequencing Options tool is supported by converging evidence that people often retain less information and misremember more when they receive complex information orally—especially when anxious or receiving adverse news—and that structured delivery, signposting, and comprehension checks improve recall and adherence. See Roy P. C. Kessels, *Patients' Memory for Medical Information*, 96 J. ROYAL SOC'Y MED. 219, 220-21 (2003) (reviewing evidence that large portions of spoken information are forgotten quickly; describing value of explicit structure and reinforcement); Victoria Siegrist et al., *Does Information Structuring Improve Recall of Discharge Information? A Cluster Randomized Clinical Trial*, PLOS ONE, Oct. 18, 2021, at 1, 10-16 (finding improved immediate recall and improved adherence after structured/signposted discharge communication); Mandhkani Mahajan et al., *The Impact of Teach-Back on Patient Recall and Understanding of Discharge Information in the Emergency Department: The Emergency Teach-Back (EM-TeBa) Study*, 13 INT'L J. EMERGENCY MED. 49, 53-55 (2020) (finding large reductions in comprehension deficits with teach-back and reporting feasibility/time burden); Philip W.B. Watson & Brian McKinstry, *A Systematic Review of Interventions to Improve Recall of Medical Advice in Healthcare Consultations*, 102 J. ROYAL SOC'Y MED. 235, 238-41 (2009) (summarizing recall-improvement interventions including repetition, summarization, and written/audio takeaways); Minh Hao Nguyen et al., *Fear and Forget: How Anxiety Impacts Information Recall in Newly Diagnosed Cancer Patients Visiting a Fast-Track Clinic*, 58 ACTA ONCOLOGICA 182, 184-86 (2019) (finding anxiety associated with reduced recall after “bad news” consultations). The tool’s plain-language and “next steps” emphasis is supported by court-specific and legal-comprehension research showing that modest word substitutions alone may not cure comprehension problems and that structure matters. Lyn Turkstra & Joseph Wszalek, *Comprehension of Legal Language by Adults with and without Traumatic Brain Injury*, 34 J. HEAD TRAUMA REHAB. E55, E62-63, E66 (2019) (finding limited benefit from tested linguistic simplifications; linking comprehension to working memory and reading fluency); Mindlin, *supra* note 278 (finding improved comprehension for plain-language court forms); Benjamin Spivak, James R. P. Ogloff & Jonathan Clough, *Asking the Right Questions: Examining the Efficacy of Question Trails as a Method of Improving Lay Comprehension and Application of Legal Concepts*, 26 PSYCHIATRY, PSYCH. & L. 441, 445 (2019) (testing structured “question trail”/fact-based sequencing approaches to improve lay comprehension and application). Finally, the sequencing tool’s procedural-justice rationale (improving understanding, reducing escalation, and increasing acceptance even when outcomes are adverse) reflects field evidence linking perceived procedural justice to trust/legitimacy in real courtroom contexts. Hilke A. M. Grootelaar & Kees van den Bos, *How Litigants in Dutch Courtrooms Come to Trust Judges: The Role of Perceived Procedural Justice, Outcome Favorability, and Other Sociolegal Moderators*, 52 L. & SOC'Y REV. 234, 251, 257 (2018) (finding perceived procedural justice associated with trust in judges, including when outcomes are unfavorable); *Procedural Justice and the Courts*, *supra* note 17, at 27-29 (synthesizing procedural-justice evidence relevant to legitimacy and acceptance).

Disagreement Protocols²⁸² are designed to allow dissent and full discussion in appellate court panels, as well as court administrative conferences. While all these tools are evidence-based, the Constructive Disagreement Protocol is the hardest to base on case-specific evidence; after all, appellate conferences are not public meetings, so they have not been studied from the inside by social scientists. Even so, these protocols are based on research in similar high-stakes professional settings. Fourth, the Decision Hygiene Starter Kit²⁸³ translates decision-science literature on

282. See *infra* pp. 224-26. The Constructive Disagreement Protocols tool is supported by research showing that dissent and structured challenge broaden thinking and reduce premature consensus, but that contrived “devil’s advocate” role-play can be less effective than authentic dissent and can heighten perceived conflict unless protected by process. See NEMETH, *supra* note 224, at 2, 39-40 (arguing that dissent broadens thinking and disrupts unanimity pressure); Charlan Jeanne Nemeth, Keith Brown, & John Rogers, *Devil’s Advocate Versus Authentic Dissent: Stimulating Quantity and Quality*, 31 EUR. J. SOC. PSYCH. 707, 710, 716-17 (2001) (experimentally comparing authentic dissent with devil’s-advocate variants; reporting superior idea generation/quality indices for authentic dissent under tested conditions). The tool’s emphasis on making dissent “safe and expected” through roles, norms, and scaffolds reflects collective-intellectual-humility research and leadership research on psychologically safe climates. See Krumrei-Mancuso et al., *supra* note 5, at 18-23 (proposing collective-level humility as a property of groups shaped by norms, roles, and procedures; describing scaffolds such as second looks and challenge roles); Silard et al., *supra* note 153, at 2, 8, 10-11 (meta-analyzing leader-expressed humility and related outcomes; supporting the importance of leader behaviors for climate). The tool’s “timeboxed roles” and moderator-style discipline draw on high-stakes professional protocols that institutionalize critique while explicitly containing interpersonal conflict. See NAT’L AERONAUTICS & SPACE ADMIN., SOFTWARE FORMAL INSPECTIONS GUIDEBOOK (1993) (describing formal inspections with defined roles and disciplined meetings aimed at defect detection); U.S. GOV’T, A TRADECRAFT PRIMER: STRUCTURED ANALYTIC TECHNIQUES FOR IMPROVING INTELLIGENCE ANALYSIS, 17, 19, 31 (2009) (providing structured techniques—key-assumptions checks, devil’s advocacy, Team A/Team B, and red-team approaches—designed to surface alternative views). The tool also incorporates implementation caution: structured techniques can fail or backfire when fidelity is low or when the method is too complex for real-world constraints. See Mandeep K. Dhami, Ian K. Belton, & David R. Mandel, *The “Analysis of Competing Hypotheses” in Intelligence Analysis*, 33 APPLIED COGNITIVE PSYCH. 1080, 1081, 1085-86 (2019) (randomly assigning analysts to use ACH or not; reporting incomplete step adherence and mixed evidence for debiasing effects). Finally, the protocols are attentive to judicial institutional realities showing that workload and collegiality costs can suppress dissent, making short, predictable disagreement slots more feasible than expecting frequent separate opinions. *Why (and When) Judges Dissent*, *supra* note 150, at 1-6 (modeling “dissent aversion” and relating dissents to workload and other institutional factors).

283. See *infra* pp. 227-30. The Decision Hygiene Starter Kit is supported by decision-science work identifying “noise” (unwanted variability) as a pervasive source of error and recommending preventive, process-based decision hygiene practices rather than bias-by-bias diagnosis. See KAHNEMAN, SIBONY, & SUNSTEIN, *supra* note 44, at 370-74 (arguing that unwanted variability is often invisible without measurement and that decision hygiene

unwanted variability (“noise”) and common cognitive errors into low-burden habits and optional self-audit routines, with

procedures can reduce error); *see generally*, Casey, Burke, & Leben, *supra* note 63, at 76-89, 91-96 (describing predictable cognitive constraints in judging and recommending structured routines and feedback practices). The tool’s sentencing anchor and self-audit logic are supported by empirical work documenting variability and disparity in judicial decisions and by experimental work showing that irrelevant anchors and other non-merits factors can influence sentencing judgments. *See* Englich, Mussweiler, & Strack, *supra* note 44, at 188, 197 (demonstrating anchoring effects in sentencing judgments by legal experts); Andrew J. Wistrich & Jeffrey J. Rachlinski, *Implicit Bias in Judicial Decision Making*, in ENHANCING JUSTICE: REDUCING BIAS 87, 92-113 (2017) (surveying cognitive processes affecting judging and cataloging debiasing strategies including “consider-the-opposite” and auditing); Wistrich, *supra* note 69, at 865-67 (discussing practical constraints and risks of truncated procedures in high-stakes judicial decision-making). The tool’s “make noise visible” and “compare like cases” elements are supported by work measuring inter-judge and regional disparities in sentencing and analogous adjudicative settings. *See* U.S. SENTENCING COMM’N, FIFTEEN YEARS OF GUIDELINES SENTENCING: AN ASSESSMENT OF HOW WELL THE FEDERAL CRIMINAL JUSTICE SYSTEM IS ACHIEVING THE GOALS OF SENTENCING REFORM 93-94 (2004) (examining inter-judge and regional disparity and the effects of structured guidelines); Crystal S. Yang, *Have Interjudge Sentencing Disparities Increased in an Advisory Guidelines Regime? Evidence from Booker*, 89 N.Y.U. L. REV. 1268, 1274-75 (2014) (analyzing changes in inter-judge sentencing disparities after the Guidelines became advisory); Ramji-Nogales, Schoenholtz, & Schrag, *supra* note 110, at 372-76 (documenting large adjudicator-to-adjudicator variation in asylum outcomes); Uhl & Pickett, *supra* note 3, at 3-5 (quantifying noise in sentencing using detailed, higher-ecological-validity materials than typical short vignettes). The tool’s “fatigue as one factor” guardrails reflect the mixed but consequential evidence on within-session effects and decision fatigue, warranting modest safeguards without overclaiming universality. *See* Shai Danziger, Jonathan Levav, & Liora Avnaim-Pesso, *Extraneous Factors in Judicial Decisions*, 108 PROC. NAT’L ACAD. SCI. U.S. 6889, 6890-92 (2011) (reporting within-session patterns consistent with extraneous influences in parole decisions); Ravi Shroff & Konstantinos Vamvourellis, *Pretrial Release Judgments and Decision Fatigue*, 17 JUDGMENT & DECISION MAKING 1176, 1177-79 (2022) (analyzing pretrial release decisions and testing fatigue-related patterns; providing nuance on mechanisms and effect size). Finally, the tool’s “structured prompts” and “implementation discipline” themes reflect evidence that small, well-designed prompts can affect sentencing judgments and that decision-support tools can have mixed effects depending on how humans use them. Eyal Aharoni et al., *Nudges for Judges: An Experiment on the Effect of Making Sentencing Costs Explicit*, FRONTIERS PSYCH., May 2022, at 1, 2, 4 (experimentally testing cost-salience prompts for sentencing decisions); Megan T. Stevenson & Jennifer L. Doleac, *Algorithmic Risk Assessment in the Hands of Humans*, IZA INST. OF LAB. ECON., Discussion Paper No. 12853, at 36 (Dec. 2019) (showing how human use of risk tools shapes outcomes and why implementation design matters); Evan M. Lowder et al., *Effects of Pretrial Risk Assessments on Release Decisions and Misconduct Outcomes Relative to Practice as Usual*, J. CRIM. JUST., March-April 2021, at 2, 2 (evaluating risk assessment effects on release decisions and misconduct outcomes); U.S. GOV’T ACCOUNTABILITY OFF., GAO-18-37, SOCIAL SECURITY DISABILITY: ADDITIONAL MEASURES AND EVALUATION NEEDED TO ENHANCE ACCURACY AND CONSISTENCY OF HEARINGS DECISIONS 42-43 (2017) (describing consistency/quality-assurance challenges and recommending evaluation measures).

sentencing cases used as an example of a recurring, high-stakes context.

Adoption of principles and practices like these does not occur by publication of them. Courts and judges adopt tools when they are reinforced by professional norms, easy to incorporate into existing routines, and visibly supported by leadership. The most straightforward pathway is judicial education. Courts can introduce Intellectual Humility by Design as a short module: a brief rationale (linking fairness, legitimacy, and decision quality), a live demonstration of two or three prompts, a one-week “try it” experiment, and a debrief. The aim is to cultivate a few repeatable habits that reduce avoidable error and improve the experience of those who come to court.

A second pathway is leadership endorsement and modeling. When chief judges and presiding judges treat intellectual humility practices as part of professional excellence—rather than as a peripheral add-on—adoption becomes legitimate. That endorsement need not be heavy-handed. It can begin with pilots and voluntary participation: a small group of judges agrees to try one or two Intellectual Humility by Design prompts for a short period and then shares what was useful, what felt awkward, and what should be revised.²⁸⁴ Courts are institutionally conservative; change is typically slow, but change can be accelerated when a new practice is framed as a refinement of the court’s core mission rather than as a cultural revolution.²⁸⁵

A third pathway is integration into existing infrastructure: bench books, courtroom scripts used by staff, standard hearing checklists, mentoring programs for new judges, and (for appellate courts) opinion templates and internal conference norms. Opinion-writing guidance has long emphasized that opinions serve multiple functions, including justifying the outcome to the parties—especially the losing party—and communicating reasons in a way that can be understood by readers beyond the immediate

284. This author would greatly appreciate that feedback.

285. *Procedural Fairness in a Pandemic*, *supra* note 154, at 702-03 (noting institutional conservatism and how external shocks can accelerate change).

legal audience.²⁸⁶ Tools that prompt a brief “letter-to-the-loser” paragraph, careful tone discipline, and clear sequencing of rule and application reflect that tradition and can be embedded in chambers practice without requiring new institutional machinery.²⁸⁷

Finally, courts can accelerate adoption through peer observation and feedback loops. Procedural fairness has already generated practical observation instruments and surveys,²⁸⁸ and those tools can also provide a platform for Intellectual Humility by Design adoption. The aim is simple: skill-building. Judges—like other professionals—improve with objective feedback, and courts can develop non-threatening ways to observe whether a judge’s practices made litigants feel heard, respected, and able to understand what happened and why.

B. Evaluation and Measurement

A design-based approach implies a measurement-based approach. If a court’s goal is to embed intellectual humility practices as routine, it must be able to detect whether those practices are present and whether they matter. The procedural-fairness literature has emphasized this point bluntly: “what we measure” shapes what we do.²⁸⁹ Yet courts often measure what is easiest—decision output and timeliness—while leaving fairness and understanding unmeasured.²⁹⁰ If Intellectual Humility by Design is to be more than a set of aspirations, courts should treat it as a set of testable interventions.

286. GARNER, *supra* note 168, at § 26.1(a) (stating that opinions serve, among other purposes, to justify the outcome “especially to the losing litigant” and that trial judges must “make the record” for review).

287. *See id.* at § 26.2(a), § 26.5(c) (noting multiple audiences and emphasizing plain, comprehensible writing and tone).

288. Rob Robinson & Danieli Evans, *Does Procedural Fairness Impact Public Perception of Judicial Opinions? Evidence from a Survey Experiment*, L. & SOC. INQUIRY, Sept. 26, 2025, at 1, 2-3.

289. *Procedural Fairness in a Pandemic*, *supra* note 154, at 700 (noting that measurement shapes performance and accountability, citing management expert Peter Drucker).

290. *Evolution of Trial Judges*, *supra* note 257, at 410-11 (noting courts’ tendency to measure efficiency and timeliness but not fairness).

Fortunately, courts do not need to invent measurement from scratch. The National Center for State Court's Court Tools Measure 1 (Access & Fairness)²⁹¹ provides a current, standardized court-user survey that has been tested for clarity and is designed to capture experiences in both in-person and remote proceedings. Two good courtroom-observation measures are available. The Center for Justice Innovation (formerly the Center for Court Innovation) has one set of measures,²⁹² and the Utah Judicial Performance Evaluation Commission has another.²⁹³ No tool is perfect, but these are feasible, familiar, and adaptable. They provide a baseline for assessing whether court users felt heard, respected, and treated fairly—and whether they understood court procedures and decisions.

Ideally, researchers can add intellectual humility in courts to their to-do lists. This might begin with a minimum viable set of measures that respects judicial workload. One component should be user perception: short surveys administered immediately after hearings or shortly after disposition. A second component should be observation: a brief checklist that tracks a small number of visible behaviors (for example, whether the judge invited each side's narrative voice; whether the judge summarized each side's key points; whether the judge stated the governing standard before announcing the outcome; and whether the explanation made clear what facts and legal rules controlled). A third component should be comprehension: a small set of questions asking whether participants understood the outcome and the reasons for it.

Courts may also wish to track modest "system" outcomes—carefully and with humility about causation. In some settings, it

291. NAT. CTR. FOR STATE CTS., COURTTOOLS MEASURE 1: ACCESS AND FAIRNESS 2 (2022), [<https://perma.cc/EZ7D-8X8R>].

292. EMILY GOLD LAGRATTA & ELISE JENSEN, CTR. FOR CT. INNOVATION, MEASURING PERCEPTIONS OF FAIRNESS: AN EVALUATION TOOLKIT (2015), [<https://perma.cc/LK6P-9D2B>]; SUZANNE SMITH, CTR. FOR JUST. INNOVATION, PROCEDURAL FAIRNESS AT COURT: A TOOLKIT FOR PRACTITIONERS, CTR. FOR CT. INNOVATION [<https://perma.cc/LC8M-F4L9>].

293. UTAH JUD. PERFORMANCE EVALUATION COMM'N, SAMPLE COURTROOM OBSERVATION REPORT 1-6 (2016), [<https://perma.cc/PLJ5-AH5X>]. Utah's six-page observation report form includes quotes about procedural justice from Burke, Leben, and Tyler. *Id.*

may be feasible to track compliance-related indicators (e.g., missed hearings, repeated hearings driven by confusion, or follow-up contacts seeking explanation). In appellate settings, it may be feasible to track a limited set of quality proxies: reversals or remands based on procedural omissions, recurring clarity problems identified by panels, or—where appropriate—plain-language readability and structure choices. None of these metrics proves that intellectual-humility practices “caused” improvement, but they can identify whether the tools correlate with outcomes courts care about and whether further refinement is warranted.

Evaluation design should fit court realities. Not every court can run a randomized experiment.²⁹⁴ But courts can do useful work with pragmatic designs: a pre/post pilot in one courthouse or docket, a phased rollout across divisions (testing the design during partial implementation), or a matched comparison between participating and non-participating judges or calendars. Qualitative feedback should complement quantitative measures: brief interviews or focus groups with litigants, lawyers, and staff can identify whether scripts felt sincere, whether time pressures were handled transparently, and whether the tools produced unintended effects (such as parties feeling rushed, patronized, or confused).

Finally, evaluation should track fidelity and adaptation. These tools are not meant to be used as rigid compliance scripts. Courts should expect local tailoring by docket type and judicial style. But evaluation must still answer a basic question: are judges actually using the core behaviors that the tool was designed to promote? If not, the tool should be revised or the adoption strategy adjusted. In a design framework, the measure of success is iterative improvement—not perfection on the first draft.

294. Hennepin County, Minnesota, once did that, though. ECKBERG & PODKOPACZ, *supra* note 191, at 12 (explaining that whether judges provided detailed on-the-bench explanations in domestic-violence cases was determined randomly, with the judge not knowing until a decision had been made which method of presenting the decision would be used).

C. Limits, Objections, and a Research Agenda

Before concluding this Article, limits and objections should be considered. In addition, since there are no reported studies of intellectual humility in a court context, research possibilities should be noted.

As for limits, this Article proposes the use of tools to increase the use of intellectual-humility principles alongside procedural-justice and decision-science lessons. Tools can be helpful, but they do not replace adequate staffing, reasonable caseloads, or competent counsel; tools will not prevent all error. But when, as is the case here, the tools are based on existing research that seems applicable to courts, courts can reasonably expect them to reduce avoidable mistakes, to improve how people experience court proceedings, and to make judicial decision-making more transparent and disciplined.

Perhaps the first predictable objection is that the inclusion of humility in the justice system is another name for deference or restraint. But intellectual humility is not the same as timidity or institutional deference. As professor Chad Oldfather explains, discussions of judicial humility often are about the “output side” of the judicial process: if the judiciary yields to other branches, it shows humility.²⁹⁵ But intellectual humility is looking at making improvements on the input side. It is a disciplined awareness of fallibility and an active willingness to test one’s conclusions against the best version of the opposing argument and the limits of the record. Properly framed, Intellectual Humility by Design strengthens neutrality and legitimacy; it does not weaken them.

A second objection is that these practices could slow courts down. In some dockets, time costs are real. That is why these tools are designed as seconds-long prompts and scripts. In addition, greater clarity on the front end can save time that might be spent later managing confusion, dealing with compliance problems, or facing escalating conflict. Still, courts need not assume time neutrality. They can test time impact in pilots and revise any prompts that create bottlenecks. A tool that adds consistent delay without measurable benefit should be shortened or abandoned.

295. See Oldfather, *supra* note 7, at 186-88.

A third objection is the risk of performative humility—scripts that sound courteous but are not experienced as sincere. That risk is genuine if judges who use these methods do not truly believe in them. Scripts are scaffolding; they can help a judge develop habits of listening, summarizing, and explaining, for example. But if used mechanically, they can backfire. Evaluation should include the perception of those who experience the court: did the parties feel heard and respected, or did the process feel dismissive? If performative humility is detected, though, the response should not be to abandon the enterprise. It would suggest a need to improve training, revise script language, and emphasize accurate summarizing and transparent reasoning as core behaviors.

A fourth objection is strategic behavior: parties may exploit openings designed to provide voice and understanding. Courts already manage risks like this through time limits, relevance rules, and firm but respectful control of court proceedings. Intellectual Humility by Design doesn't require unlimited voice; it requires meaningful opportunity to be heard and an accurate demonstration that the judge understood the key points. Done well, it can reduce gamesmanship by making the decision's reasons and constraints transparent.

Finally, one might object that humility practices—especially those that include acknowledging human impact—threaten neutrality or resemble empathy-based favoritism. Again, the framing matters. Respectful acknowledgment is not a promise to rule for a party; it is a recognition that the case matters to the people involved.²⁹⁶ Reasons-first transparency is often the best neutrality signal a court can give. When courts explain the governing rule and the controlling facts plainly, they strengthen legitimacy even when the outcome is unfavorable.

These objections also point to a research agenda. First, courts should test whether Intellectual Humility by Design benchmark practices improve procedural-fairness perceptions and comprehension, as well as whether effects differ by docket type (such as high-volume calendars, self-represented litigants, and

296. *Expectation of Empathy*, *supra* note 23, at 50 (setting forth the argument that the public wants judges to be responsive and empathetic but also neutral and fair).

civil versus criminal). Second, courts should test which adoption pathways work best: judicial education alone, leadership endorsement, mentoring and peer observation, or integrated approaches. Third, appellate courts should test whether different drafting prompts (reasons-first structure, clear identification of a losing party's best points, tone discipline, and modesty about the scope of a decision) improve understanding and reduce recurring clarity problems.

This agenda is not merely aspirational. Courts already have mechanisms—education programs, bench cards, observation tools, and surveys—that can be used for careful, incremental evaluation. If intellectual humility is to be by design, courts should treat these tools as a starting point, not an end point: adopt, test, revise, and re-test. That is how good institutional practice becomes durable.

VII. CONCLUSION

The rule of law depends on more than correct outcomes in individual cases. It depends on sustained public acceptance of courts as legitimate institutions—acceptance strong enough that most people will comply with orders, respect procedures they do not control, and accept decisions they may dislike. That acceptance is under pressure in an era of polarization, misinformation, and declining trust in institutions. Courts cannot repair those broader forces by themselves, but they can control something central: what people experience when they encounter the justice system.

Judges are not private professionals pursuing private ends; they are public officials exercising public power. And because their authority depends so heavily on public acceptance, judges have a special obligation to make their work intelligible to those they serve. As two judges have put it in the context of public engagement, “[w]e are public servants, not disengaged robed philosophers, and the public has a right to know who we are and what we do.”²⁹⁷ A justice system that is experienced as opaque,

297. Stephen Louis A. Dillard & Bridget Mary McCormack, *The Robed Tweeter: Two Judges' Views on Public Engagement*, 20 J. APP. PRAC. & PROCESS 179, 181-82 (2019); *Confirmation Hearings on Federal Appointments: Hearing Before the S. Comm. on the*

indifferent, or rushed will spend its legitimacy capital quickly—even when outcomes are legally correct.

That is where this Article's themes converge. Procedural-justice research shows that perceptions of legitimacy turn heavily on whether people feel heard, treated with dignity, and given transparent reasons by a neutral decision-maker.²⁹⁸ Decision-science research shows that even conscientious judges operate under cognitive constraints—fatigue, distraction, time pressure, and unwanted variability—that can degrade decision-making and make fairness harder to demonstrate.²⁹⁹ Together, these bodies of work suggest a practical institutional goal: courts should make careful decision-making *visible*—not as performance, but as routine professional practice.

Intellectual humility—disciplined awareness of fallibility and a willingness to test one's conclusions against the strongest version of the opposing position³⁰⁰—offers a bridge between legitimacy and accuracy. It is a disciplined commitment to do the work of judging in a way that is careful, transparent, and bounded: reasons-first, record-faithful, attentive to limits, and responsive to the losing party's best arguments.

The right time to begin this work is not after the next legitimacy shock, the next viral misimpression, or the next crisis of confidence. It is now. Indeed, former Michigan Chief Justice Bridget Mary McCormack has argued that judges “are ethically obligated to improve the judicial system over which they preside.”³⁰¹ This Article offers low-cost mechanisms judges and courts can use to pursue that obligation in their daily work—bench prompts, drafting templates, and adoption pathways that can be piloted, measured, and refined. If legitimacy is the operating capital of the rule of law, then intellectual humility by design offers a practical way to replenish it—case by case, opinion by opinion, and courtroom by courtroom.

Judiciary, 112th Cong. 36 (2011) (statement of Jacqueline H. Nguyen, Nominee to be U.S. Circuit Judge for the Ninth Circuit) (“The judiciary is a public institution, and judges are public servants.”).

298. *Expectation of Empathy*, *supra* note 23, at 50.

299. *See supra* note 196 and accompanying text.

300. *See supra* note 4 and accompanying text.

301. Bridget Mary McCormack, *Staying Off the Sidelines: Judges as Agents for Justice System Reform*, 131 YALE L.J.F. 175, 175 (2021).

Intellectual Humility by Design Bench Card

Recognize · Reflect · Respond · Reassess (R4)
A practical routine for trial and appellate judges

Quick-use checklist:

1. Recognize fallibility

- What would change my mind?
- Consider the opposite: write two plausible counterpoints.
- Premortem: if this decision is later criticized or reversed, why?
- Outside view: how would a neutral colleague frame this?
- Reminder: revision is progress; better reasons are the goal.
- Trial: name one uncertainty or assumption you should test before ruling.
- Appellate: note a plausible alternative reading and the limits of the holding.

2. Reflect before ruling

- Voice first: “Here is what I hear—did I get that right?”
- Ask each side (one sentence): “What is your strongest point, and the best legal authority for it?”
- Quick procedural-justice check: voice, neutrality, respectful treatment, and trustworthiness.
- Trial: Paraphrase each party’s key point; invite a one-sentence correction; pause or take the matter under advisement if tempers rise.
- Appellate: In a bench memo, write one sentence stating each side’s best case; in conference, start with that before voting; ask what you might be missing.

3. Respond with reasons

- Roadmap early (point-first when practical): issue/outcome → rule/standard → key facts → analysis → remedy.
- Address the strongest opposing point (as the party would state it).
- Name constraints (standard of review; statutory bounds; record limits).
- Add one sentence recognizing the human stakes (tone signals care, not bias).
- Trial: Give step-by-step oral reasons; make next steps clear.
- Appellate: Engage the best contrary authority respectfully; explain why your reading prevails; keep the holding bounded.

4. Reassess after decision

- Second look: read as the losing party—did I fairly state and answer the best point?
- Invite perspective; treat disagreement as data.
- Optional: 15-minute quarterly “R4 retro” (review any reversals, review decisions/orders for unclear passages; make one template or checklist tweak; log it for next R4 retro).
- Trial: Do a brief tone and clarity pass; update templates; log one improvement.
- Appellate: Ask a colleague to present the strongest contrary argument; revise where others expose gaps.

Note: A more extensive version of this bench card will be posted on the website proceduraljustice.org

How to use:

- Pick one prompt from each R when the case is routine; use more prompts when the stakes, uncertainty, or time pressure increase.
- Use “in chambers” prompts before issuing a decision; use “on the record” prompts to make voice, neutrality, and constraint visible.

In chambers (trial and appellate):

- Write the best argument for the outcome you are not inclined to reach—using the other side’s own framing.
- Identify one assumption you are making (about facts, credibility, or law) and what evidence would test it.
- Ask: “If I am wrong, where am I most likely wrong—facts, law, or procedure?”
- Use a short premortem: “A reviewing court (or the public) criticizes this—why?”
- When you disagree with a colleague’s view, treat it as information: “What are you seeing that I am not?”

On the record (trial):

- Summarize what you heard before ruling; invite a one-sentence correction.
- Explain the next step (what happens now; what the party must do; when the court will act).
- If emotion is high, name the process move: “I’m going to pause so I can decide carefully.”

Conference and drafting (appellate):

- Start conference by stating each side’s best point (even if you will reject it).
- Assign a rotating “counterargument” role: one judge must present the strongest contrary authority and the best response.

- During drafting, mark the holding and its limits; avoid unnecessary breadth.
- Do a “losing party” read for tone and clarity; revise to address the best point directly.

This appendix tool is intended as a practical translation of the article’s evidence base; supporting authorities appear in the article text and footnotes.

Reasons and Explanations Toolkit
(Plain-Language Layer + Reasons Discipline)

Goal: Make reasons do double duty—improving decision quality and making fairness observable, especially to a losing party.

A. Overriding requirements for any ruling or opinion

- Orient the reader/listener early: what decision is being made and what questions must be answered.
- Engage the strongest opposing point in its best form (avoid responding to a weaker version).
- Make the decision path transparent: rule/standard → key facts → application → conclusion.
- Include a plain-language layer (especially in trial rulings and for self-represented court users).
- State constraints and limits: what the court can and cannot decide; what is and is not being decided.
- Provide next steps (trial courts) or clarify what follows from the holding (appellate).
- Use respectful acknowledgment: recognize what was at stake without signaling partiality.

B. Plain-Language Layer (5-7 sentences)

Use this as the “closing summary” in an oral ruling or as a short segment in a written decision. A short plain-language layer can materially improve nonlawyers’ understanding of what the court decided and why. (Number of sentences is suggestive only.)

- What is being decided today (one sentence).
- The rule or standard that controls (one sentence).
- The key fact(s) that mattered most (one sentence).
- The court’s conclusion (one sentence).
- The strongest opposing point and why it does not change the result (one or two sentences).
- Scope/limits (optional one sentence).
- Next steps (trial) or effect of the holding (appellate) (one sentence).

Two model examples (adaptable)**1. Trial oral ruling (template):**

- “The question today is _____.
- The law requires me to decide it by applying _____.
- The facts I find most important are _____.
- For those reasons, I am ordering/denying _____.
- I have considered _____ (the strongest point you raised), but it does not change the outcome because _____.
- This ruling decides _____ and does not decide _____.
- The next steps are _____.”

2. Appellate opinion (simplified checklist):

- We decide whether _____.
- Applying _____ (standard), we hold _____.
- The determinative facts are _____.
- The appellant’s strongest contrary argument is _____, but it fails because _____.
- Our holding is limited to _____.

C. Reasons Discipline Checklist

Use this as a pre-issuance “second look” for a solo judge or as a final circulation check for a panel.

1. Quick-use (30-60 seconds)

- Have I identified the real question and the decision standard?
- Have I answered the strongest opposing point (not a straw-man)?
- Can I state the determinative facts in two sentences?
- Did I explain the step that most changes the outcome (the decisive point)?

- Should I add a short scope/limits sentence to prevent overread?
- For trial rulings: did I add next steps in plain language?

2. *In chambers / drafting (3-5 minutes)*

- Counterargument test: write the best argument for the other result in one paragraph. Did I respond to that argument?
- Constraint test: name the strongest legal authority or factual constraint that limits discretion here. Did I follow and state it?
- Noise check: if another judge saw the same record, what is the most likely place we would differ—and why?
- Legitimacy check: if I were the losing party, what would I say the court did not understand? Did I answer that?
- Plain-language check: can a non-lawyer restate the ruling’s reason without guessing?

D. Cautions

- Avoid “performative complexity.” The goal is clarity and fair engagement, not length.
- Avoid “false balance.” Engaging the strongest opposing point does not require treating weak points as equal.
- Don’t confuse respectful acknowledgment with agreement. Respect is a process cue; the law controls the outcome.

Trial Oral-Ruling Sequencing Options

Goals: Improve understanding, reduce unnecessary escalation, and make fairness observable—without adding needless length.

A. Default principles (use with any sequence)

- Signpost your structure (one sentence). “I’m going to do three things: (1) my decision, (2) the reasons, and (3) what happens next.”
- Orient first. State what is being decided today and what question(s) the court must answer.
- Use one closed-loop listening move (10-20 seconds). “Let me make sure I understood your key point: _____. Did I get that right? If not, what would you add in one sentence?”
- Keep reasons to the decision-controlling points. Use two (or three) decision-controlling reasons; avoid narrating everything.
- End with a plain-language close. What I decided, why (two reasons), and what happens next.
- If emotion spikes, name the process move. “I’m going to pause so I can decide carefully.” If needed, take the matter under advisement rather than escalating.

B. Quick checklist (10 seconds before you end)

- Did I state the question, the standard, two hinge reasons, the strongest opposing point, the decision, limits, and next steps?
- Did I do one comprehension check (understanding check—called teach-back in healthcare—or brief deadline restatement)?

C. Choose a sequencing style

1. Option A — Path-first

- Issue → Standard → Key facts → Decision → Next steps

- Use when: self-represented parties; high emotion; or you suspect people will stop processing once they hear the result.
- Cautions: keep it tight; don't start with too much background.

Script (adaptable):

- "The question today is ____."
- "The law requires me to apply ____."
- "The facts I find most important are ____."
- "For those reasons, I am ____."
- "Next steps are ____."

2. Option B — Result-first with a promise (a "retention device")

- Result → Promise reasons/next steps → Issue framing → Reasons → Next steps → Plain-language close
- Use when: immediate clarity matters (release conditions, scheduling, immediate relief) or the result itself reduces uncertainty.
- Cautions: don't delay the explanation after announcing the result.

Script (adaptable):

- "My ruling is _____. I will explain the reasons and the next steps."
- "The question I must answer is _____."
- "The controlling rule is _____. Two facts matter most: _____ and _____."
- "That is why the result is _____."
- "Next steps: _____."

3. Option C — Two-pass hybrid (best for adverse/high-conflict/high-stakes)

- Forecast → Bottom line → Pause → Summarize reasons (2-3) → Summarize next steps + key deadline → Brief structured reasons → Recap

- Use when: you're delivering bad news; emotions are high; obligations and deadlines are complex.

Script (adaptable):

- Forecast: "This may be difficult to hear."
- Bottom line: "My ruling is ____."
- Pause (3-8 seconds).
- Summarize reasons (2;3 bullets): "Two reasons control: ____; ____."
- Summarize next steps: "Next: ____ by ____."
- Brief structured reasons: "On the first question— ____—I find _____. On the second— ____—I find ____."
- Plain-language close: "To summarize: the decision is _____. Two reasons are ____ and _____. Next steps are ____ and ____."

D. Retention devices (use especially when the outcome is adverse or the order is complex)

1. *Headline recap (10-20 seconds)*

- "To summarize: the decision is _____. Two reasons control: ____ and _____. The next step is ____ by ____."

2. *Written or replayable takeaway*

- "You will receive a written order today that lists what you must do, deadlines, and how to seek review."
- In remote hearings: put the same bullets in the chat or send a same-day message that mirrors your oral structure.

3. *Understanding check (one question; carefully framed)*

- "So I can be sure I explained it clearly: tell me, in your own words, what you believe you need to do next and by what date."
- If incorrect: "Let me restate that as clearly as possible . . ."

4. *Bounded Q&A (1-3 minutes; preserves voice without retrying the case)*

- “What questions do you have about what the order requires or the timeline?”
- “I can clarify what the order says and what the deadline is. I can’t give legal advice or re-argue the evidence.”

E. Situational modules

1. *High-volume docket (≤90 seconds)*

- Use Option A: Decision → 2 reasons → 1 deadline, plus a brief understanding check: “Tell me the next date you must remember.”

2. *Remote or virtual hearings*

- Increase signposting and deliberate pausing (turn-taking is harder).
- Provide an immediate written recap that duplicates the oral recap.

3. *Language access/low literacy/cognitive impairment*

- Confirm interpreter needs before ruling.
- One obligation per sentence; concrete verbs; avoid relying on minor word substitutions alone.
- Provide a written plain-language summary and use an understanding check.

4. *Volatile situations/safety concerns*

- Prioritize safety: shorten, slow down, and consider under-advisement if explanation would inflame risk.
- Still deliver next steps and restrictions clearly; confirm understanding.

**Constructive Disagreement Protocols
(Conference and Administrative Meetings)**

Goals:

- Improve decisions and written products by surfacing decision-controlling concerns early.
- Protect collegiality by keeping critique about reasons, assumptions, and risks—not people.
- Add structure without materially extending meeting time (default: 2-5 minutes per matter).
- Reduce “late issues” raised after circulation by creating a predictable, safe moment to speak up.

Note: These protocols draw on evidence from other high-stakes professional teams and on what we know about judicial dissent incentives. Because appellate conferences are private and vary across courts, treat this as a menu of time-bounded options to pilot and adapt.

A. Quick-use checklist (30 seconds)

- Name the question and the tentative disposition.
- Surface one decision-controlling counterargument (even if you expect unanimity).
- Identify one assumption that “must be true” for the reasoning to hold.
- Decide whether the issue is routine (brief) or hard/high-salience (full).
- Close with: “Any concerns that would strengthen the opinion or the plan?”

B. Default protocol for routine matters (brief, 2 minutes)

1. Disposition in one sentence (by the presiding judge or case presenter).
2. One-round “issue-spot” prompt: each judge offers either (a) “no concern” or (b) one sentence naming the strongest decision-controlling concern.
3. If a concern is raised: decide whether it can be handled now (quick fix) or should trigger the “full” protocol.

Note: If your court speaks in seniority order, keep that norm. The short-form (brief) protocol can be done within that order by adding a single, predictable “one concern” sentence per judge.

C. Full protocol for hard or high-salience matters (5 minutes)

1. *Independent-first tentative position (30-60 seconds)*

Before discussion, each judge writes two short lines: (a) tentative disposition; and (b) one reason you might be wrong. This protects independence and makes it easier to raise concerns without personal friction.

2. *Seniority-order discussion (keep the court’s norm)*

If your court speaks in seniority order, do not fight it. Use the order, but add a structured “challenge point” at the end.

3. *Strongest opposing point round (time-limited, 2 minutes total)*

- a. One judge states the strongest decision-controlling counterargument (60-90 seconds). Prefer a judge who genuinely has doubts or is least certain.
- b. The tentative majority answers in one minute: “Here is our best response.”
- c. Close with a single question: “What would you change in the reasoning to address that point?”

4. *Key assumptions check (optional, 1 minute)*

Name 2-3 assumptions that must be true for the result to hold. For each: “What could break this?” If an assumption is shaky, assign a quick verification step before finalizing.

5. *Micro premortem (optional, 1-2 minutes)*

Prompt: “It’s two years later and this decision created serious problems or was reversed. Why?” Round-robin: one reason per judge; capture the top two risks to mitigate in the opinion.

D. Protecting collegiality (rules of engagement)

- Critique reasons and assumptions, not motives: “I’m worried about X,” not “You’re wrong.”
- Use neutral language: “The strongest opposing point is . . .”
- No interruptions during the opposing-point round.
- If a concern is raised, the question is “How do we strengthen the reasoning?”
- If tension rises, pause and restate the shared aim: accuracy, fairness, and institutional credibility.

E. Preventing “late issues” after circulation

- Default expectation: decision-controlling concerns should be raised in conference when possible.
- If a new concern emerges during drafting review, raise it promptly and pair it with a proposed fix (one sentence) to minimize disruption.
- If the concern would change the result or require major restructuring, consider re-conferencing the panel (time-limited) rather than trading escalating memos.

F. Trial-court administrative meetings (adaptation)

Use the same logic for policy and operations decisions: build a brief, predictable challenge point into the agenda.

- Before voting, add a 2-minute “risk check”: “What could go wrong, and who will be affected?”
- Rotate a “risk raiser” role only for agenda items where no one has voiced a concern (to avoid performative dissent).
- Timebox: 60 seconds to state the strongest concern; 60 seconds to state the best response; then decide next step (adopt, revise, pilot).
- Close with an implementation check: “What will we measure to know if this worked?”

Decision Hygiene Starter Kit (Trial and Appellate)

Note: These options draw on research about noise (unwanted variability), cognitive bias, and high-stakes professional decision-making. Courts and case types vary, so treat this as a menu of time-bounded practices to pilot and adapt.

Goals:

- Reduce unwanted variability across similar cases by using consistent steps and prompts.
- Make the reasoning process more reliable without turning every decision into a research project.
- Catch mistakes early (before finalizing) and reduce late-stage revisions.
- Support confidence and legitimacy by showing disciplined, evenhanded decision-making.

A. Quick-use checklist (60-90 seconds)

- State the decision and the standard in one sentence.
- Name the two or three decision-controlling facts.
- Run one “opposing-view check”: what is the strongest reason the other side should win?
- Confirm comparators: what similar cases am I implicitly comparing to (if any)?
- Confirm the next step: written order language, conditions, or deadlines.

B. Sentencing workflow (anchor protocol)**1. Pre-hearing preparation (2-5 minutes)**

- a. Write a tentative range/outcome and one sentence why (for yourself).
- b. List two case-specific factors that could justify a different outcome (mitigating and aggravating).
- c. Identify one “anchor risk”: a number or recommendation that might pull the sentence (charging decision, plea offer, guideline midpoint, probation recommendation).

- d. Decide what you should hear at the hearing to confirm or change your tentative view.

2. In-hearing discipline (30-60 seconds)

- a. Let each side state the single best point they want you to remember (one sentence).
- b. Ask one clarifying question tied to the standard (not general narrative).
- c. If you feel pulled by a number, pause and restate the governing reasons before selecting the outcome.

3. Before announcing the sentence (60-90 seconds)

- a. Two-sentence reasons: (a) the standard; (b) the decision-controlling considerations.
- b. Symmetry check: would I say the same thing if the parties were reversed or the person were different?
- c. If the case is close, run a micro premortem: "If this is later viewed as wrong or unfair, why?" (one reason).
- d. State next steps (conditions, deadlines, credit time, review rights) clearly.

C. Self-audit (lightweight, optional but recommended)

The point is not to "grade" yourself. It is to make noise visible and catch drift early.

1. What to track (pick one)

- Sentencing outcomes for a single recurring offense category (with a few basics: history score and key aggravators/mitigators).
- Bail/pretrial release decisions for one recurring category (if you handle them).
- A recurring civil ruling category (continuances, discovery sanctions, protective orders).

2. *A simple monthly review (15-30 minutes)*

- a. Look at 10-20 recent decisions in the category (or fewer if volume is low).
- b. Ask: are outcomes clustered, or are there unexplained outliers?

3. *If you see drift, identify a “reason pattern” (what changed?) and one guardrail to restore consistency.*

D. Built-in “second look” for high-salience cases (2-3 minutes)

- Write the strongest opposing argument in two sentences.
- Write your best answer in two sentences.
- Identify one fact you might be overweighting and one you might be underweighting.
- If still uncertain, narrow the decision: what is the smallest change that would address the concern?

E. Fatigue and pacing guardrails (one factor; basic safeguards)

- If possible, avoid stacking the most consequential sentencings at the very end of a long calendar.
- Use a short reset when you notice speed-up: stand, breathe, sip water, or take 60 seconds before ruling.
- When time is tight, shorten everything except the decision-controlling reasons and the next steps.

F. Appellate adaptations (conference and circulation)

1. *Conference micro-structure (2-5 minutes per case)*

- a. Each judge comes with a tentative disposition and one reason they might be wrong.
- b. Use the court’s normal speaking order (including seniority order).
- c. Add a timeboxed “strongest opposing point” round: 60-90 seconds to state it; 60 seconds for the response.
- d. Identify one verification task before assigning final writing (recorded in each judge’s notes).

2. Circulation guardrail (prevent late surprises)

- a. When reviewing a draft, raise any decision-controlling concern early and pair it with a proposed fix (one sentence).
- b. If the concern would change the result, consider a brief re-conference rather than escalating memo exchanges.

**Decision Hygiene Self-Audit Log
(Sentencing Anchor)**

Purpose: Personal consistency check (not evaluation). Pick one recurring category.

Instructions: Fill 10 to 20 rows per month (or fewer if volume is low). Review monthly for outliers/drift.

Case No.	Date	Category	History	Outcome	Key reasons (1 sentence)	Outlier? why
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

Case No.	Anchor risk (Y/N)	Aggravator (1)	Mitigator (1)	Comparator note (optional)
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

Monthly 3-question review:

1. Clustered or scattered outcomes for similar profiles?
2. Any unexplained outliers or surprises?
3. One guardrail to try next month: one prompt, one pause point, or one comparators check.