IS JUDICIAL EXPERTISE DYNAMIC? JUDICIAL EXPERTISE, COMPLEX NETWORKS, AND LEGAL POLICY

Anne Lippert[*] & Justin Wedeking[**]

TABLE OF CONTENTS

*   University Research Postdoctoral Fellow, Departments of Political Science and Psychology, University of Kentucky, amli237@uky.edu.
**   Associate Professor, Department of Political Science, University of Kentucky, justin.wedeking@uky.edu, (859) 257-7040.

INTRODUCTION

In recent years, legal scholars have debated whether the law and the United States legal system are becoming more complex.[1] One common theme throughout this body of research is the general argument and conclusion that complexity is indeed growing. In fact, Schuck even argues that the growth thesis is fairly uncontroversial.[2] But what is generally meant by this growing complexity? To start, we adopt Posner's framework, which argues that there are two related types of complexity.[3] The first is external complexity, which is defined as complexity due to the external environment.[4] The external environment is composed of such systems from the economic, political, ecological, or technological arenas and the interactions within and between them.[5] By almost all accounts, the external world is becoming more complex, and this, according to Posner, makes the job of judging more difficult, especially because judges have no control or influence over the external environment.[6]

In contrast, as Posner argues, internal complexity is brought on by judges' attempts to manage and cope with the rising levels of external complexity, usually by employing a complex style of legal analysis to resolve cases without having to understand factual complexities.[7] The important intersection between internal and external complexity suggests that judges may either "complexify" the law and legal process needlessly, or they may try to simplify the law at the expense of the complex external world.[8] Hence, increasing internal complexity only adds to the difficulties that judges face in carrying out the duties of their job.

---

1.   *See generally* Michael J. Bommarito II & Daniel M. Katz, *A Mathematical Approach to the Study of the United States Code*, 389 PHYSICA A 4195 (2010); *see also* RICHARD A. POSNER, REFLECTIONS ON JUDGING (2013); Peter H. Schuck, *Legal Complexity: Some Causes, Consequences, and Cures*, 42 DUKE L.J. 1 (1992).
2.   Schuck, *supra* note 1, at 6.
3.   POSNER, *supra* note 1, at 3-4.
4.   *Id.* at 4.
5.   *Id.* at 3-4.
6.   *Id.*
7.   *Id.* at 4.
8.   *Id.*

Thus far, claims of growing complexity have received relatively little challenge. As a result, it seems implicitly assumed that judges would develop expertise through experience that would enable them to cope with this growing complexity. But why should we care about growing complexity and whether judges develop expertise to cope with it? To start, the implications are enormous. Generally speaking, there are potentially severe negative consequences for our law and legal system that stem from growing complexity. Specifically, growing complexity and a lack of expertise may have unhealthy effects on U.S. democracy if they hinder the ability of judges to render just and fair decisions; if they result in barriers for litigants to gain access to the legal system; and, more generally, if they impede our society's system of governance that was designed to equitably and efficiently resolve disputes. In short, the growing complexity of our legal system and a lack of judicial expertise present an important concern.

More specific to our purposes here, for those who acknowledge that the legitimacy of a legal system is necessary for a thriving and healthy democracy, how judges cope with this growing complexity becomes vitally important. Interestingly, however, what is generally missing from this discussion on complexity is a systematic examination of how judges have dealt with the growing complexity of the legal system. Hence, the broader question we address is, how have United States Supreme Court Justices' expertise fared despite this growing complexity? This question is necessary for two different yet related reasons.

First, as the external environment changes and becomes more complex, a person's expertise can change *even if the person does not change their approach*. Stated differently, a person's expertise may actually *decline* if their method of coping with the increased complexity is not sufficient. Second, and by implication, in order for an individual to *increase* or develop his or her expertise in an environment that is becoming more complex, an individual must adapt. This essentially means that judicial expertise needs to account for how individuals (judges mainly) themselves grapple with a changing external world. This often comes in the form of a struggle to incorporate increased amounts of information and knowledge into legal decision-making and opinion-writing.

We argue that a dynamic understanding of expertise is possible when we conceptualize expertise in a way that is consistent with research from psychology on domain expertise. Specifically, expertise is a function of two things. First, expertise consists of how

much knowledge one possesses and displays in domain relevant tasks and decisions.[9] Second, experts and novices differ in how knowledge is structured and stored or "chunked."[10] Thus, if we are to understand the importance of judicial expertise and how it operates within an environment of growing complexity to influence the law and our legal system, we first need to measure the Justices' expertise in a dynamic manner. By first developing a new measure of judicial expertise, we can address whether the Justices' expertise has increased or decreased during this changing external environment.

To first develop a new measure of expertise, we examine the written majority opinions from twenty-seven Supreme Court Justices who all served at least ten terms over the last century. Examining their written opinions from two issue areas—criminal procedure and economic activity—we adopt a "text as data" approach and treat each written opinion as a particular type of complex network— namely, a co-occurrence network—where each word in the text represents a node, and nodes are connected by a link if their corresponding words occur next to each other (reading left to right) in text. Representing opinions as networks allows us to analyze the topological features of the obtained representation and extract a set of informative network measurements that are indicative of the author's expertise on the two established dimensions.

We use these indicators to develop estimates of judicial expertise corresponding to each year. Using this new, dynamic measure, we provide a descriptive account of how Supreme Court Justices' expertise has evolved over time in their attempts to grapple with a changing information environment that is becoming more complex. We find that some Justices increase their expertise over time, even in the face of growing complexity of the external environment. However, we also find that some Justices' expertise has not kept pace and has, in some cases, declined. This suggests to

---

9. *E.g.*, K. Anders Ericsson, Ralf Th. Krampe & Clemens Tesch-Römer, *The Role of Deliberate Practice in the Acquisition of Expert Performance*, 100 PSYCHOL. REV. 363, 365 (1993); Ronald T. Kellogg, *Professional Writing Expertise*, *in* THE CAMBRIDGE HANDBOOK OF EXPERTISE AND EXPERT PERFORMANCE 389, 389 (K. Anders Ericsson et al. eds., 2006).

10. *E.g.*, Dennis E. Egan & Barry J. Schwartz, *Chunking in Recall of Symbolic Drawings*, 7 MEMORY & COGNITION 149, 149 (1979); K. Anders Ericsson, William G. Chase & Steve Faloon, *Acquisition of a Memory Skill*, 208 SCIENCE 1181, 1181 (1980); K. Anders Ericsson & Peter G. Polson, *An Experimental Analysis of the Mechanisms of a Memory Skill*, 14 J. EXPERIMENTAL PSYCHOL.: LEARNING, MEMORY, & COGNITION 305, 305-06 (1988); William G. Chase & Herbert A. Simon, *Perception in Chess*, 4 COGNITIVE PSYCHOL. 55, 55-56 (1973).

us that expertise is not simply a function of experience. Rather, the acquisition of legal expertise is complicated because of the rapidly changing external environment and how Justices have chosen to cope with it.

This paper makes several contributions. First, this is the first paper to develop a measure of expertise of the Justices that is based on the content of their work product—their written opinions. Second, it is the first work to use networks derived from text to differentiate between novices and experts in multiple issue areas. Third, we show that legal expertise is structured by two factors—the amount of knowledge of the expert and how that knowledge is structured—and that these two factors are consistent across two issue areas. Fourth, this paper illustrates the dynamic nature of expertise in an unstable domain—with Justices showing both increased and decreased expertise. This provides a more nuanced understanding of how expertise can evolve in the face of an increasingly complex environment. In our conclusion, we highlight future directions for research in the area of expertise.

The rest of this Article proceeds in the following fashion. In the next section, we add more detail to our understanding of Posner's two types of complexity and its connection to expertise. From there, we discuss how previous research has typically conceptualized expertise, drawing on research in the fields of political science and then psychology. Following that, we outline our methodological approach to generating a new, dynamic measure that incorporates network indicators derived from texts of Supreme Court opinions. We then examine trends of the Justices' expertise over time, and also whether greater expertise correlates with a decrease in the number of concurring opinions. We conclude with a brief discussion of future directions for our research as well as some implications of our work.

I. GROWING COMPLEXITY OF THE LEGAL SYSTEM

Research on the Supreme Court illustrates examples of both types of Posner's complexity. For example, the growth over time of *amicus curiae* (friend of the Court) brief submissions illustrates the growing complexity of the external environment. Amicus briefs are often thought of as sources of extra information for the Justices about the case that might not be included in the parties' briefs, yet could be vitally relevant to the case. Collins examines both the total number of briefs submitted each term, as well as the total number of participating outside parties each term, and shows how they have

changed over time.[11] Specifically, Collins finds that both of them increase sharply over time.[12] If increased amicus participation is taken as a sign of the complexity of a case, which seems like a fair assumption given that amicus briefs often add much nuance and varied perspectives to a case, then the amount of information being communicated to the Court has increased substantially and suggests that we have a more complex external legal environment compared to earlier years. In other words, the external environment brings cases to the doorstep of the Supreme Court that are now more complex than past cases. This only further bolsters the need to know whether Justices are developing the expertise needed to keep pace.

Turning to an example of internal complexity, scholars have highlighted the growing length of Supreme Court opinions. In attempting to communicate with their audiences, Justices have the choice of how much or little they want to write when drafting opinions. Black and Spriggs graphically display the median length of Supreme Court opinions over time, as well as the inter-quartile range over time.[13] Their analysis reveals an obviously upward trend over time.[14] In short, Supreme Court opinions are becoming dramatically longer (e.g., the median opinion length in 2000 had approximately 4,500 words while from the 1940s to the 1970s it was approximately 2,000 words).[15] If we assume that longer is generally equivalent to being more complicated, it suggests the possibility that judges are coping with the growing external complexity by writing longer opinions.[16] These increasingly lengthy opinions led to Posner's advice that judges need to write shorter opinions.[17] To demonstrate this, in the appendix of Chapter 8, he goes to great lengths to

---

        11.   *See* Paul M. Collins, Jr., *Friends of the Court: Examining the Influence of Amicus Curiae Participation in the U.S. Supreme Court Litigation*, 38 L. & SOC'Y REV. 807, 810-11 (2004).

        12.   *Id.*

        13.   Ryan C. Black & James F. Spriggs II, *An Empirical Analysis of the Length of U.S. Supreme Court Opinions*, 45 HOUS. L. REV. 621, 635 (2008).

        14.   *Id.*

        15.   *Id.*

        16.   We argue that most people will likely think this a safe assumption. We note, however, that growing the length of an opinion might also be construed, in some contexts, as a Justice's attempts at greater specificity and precision (i.e., being more thorough). That greater length may be an attempt to carefully, in detailed fashion, sort through the complicated details and articulate the many nuances of a case.

        17.   POSNER, *supra* note 1, at 236, 255-58, 276-86.

demonstrate this point by rewriting a 3,237 word appellate court decision (that is not his) to a 602 word opinion.[18]

What is unclear, however, despite Posner's prescription, is under what circumstances will shorter opinions indicate more expertise? Perhaps the most challenging aspect to his advice is that Justices will have to omit domain relevant knowledge to write shorter opinions, and that may be problematic because it can leave counterarguments unaddressed. Perhaps the key is whether adding length to an opinion is adding new or redundant knowledge, which is something we will return to later. Regardless, to better understand why we need a new, dynamic measure of judicial expertise, in the next section we review current literature on how political scientists have measured judicial expertise and illustrate how none of those approaches can accommodate a changing information environment.

II. EXPERTISE IN POLITICAL SCIENCE

Previous attempts to quantitatively measure expertise usually equate experience with expertise. Specifically, prior research has looked at expertise under three different umbrellas. The first approach has looked for what is known as "the freshman" effect, where a Justice has an acclimation period, usually lasting through the first or second term on the Court, where they are orienting themselves to their new job, role, and judicial philosophy.[19] This usually entails, among other things, simultaneously developing the skills of judging, bargaining, and negotiation with their colleagues; adjusting to their new environment; and crafting written opinions on a collegial court. In addition, Wood, Keith, Lanier, and Ogundele found significant differences for both the time period and if they were previously a lower court judge.[20] In particular, modern era judges experienced stronger acclimation effects than those from an earlier era (1888-1940), and those that lacked lower court judge experience were susceptible to greater acclimation effects (e.g., voting instability).[21] In contrast, Scheb and Ailshie find that Justice O'Connor did not exhibit behavior that was consistent with the

---

18. *Id.* at 276-86.
19. *See generally* Timothy M. Hagle, *"Freshmen Effects" for Supreme Court Justices*, 37 AM. J. POL. SCI. 1142 (1993).
20. Sandra L. Wood et al., *"Acclimation Effects" for Supreme Court Justices: A Cross-Validation, 1888-1940*, 42 AM. J. POL. SCI. 690, 690-94 (1998).
21. *Id.*

"freshman effect" label,[22] and Rubin and Melone reached a similar conclusion about Justice Scalia,[23] as well as Justice Kennedy.[24]

This approach has some empirical validity and is highly intuitive, but we note that it also more or less assumes that Justices make the shift from novice to expert after only a short amount of time with no consideration for how the external environment has changed and little systematic attention to the outputs of judging—opinions. Furthermore, after accounting for a number of complicating factors that resulted in previous studies reporting mixed results, Hagle finds that some Justices experience freshman effects while others do not.[25] This further suggests the need to find an alternative approach for the measure of legal expertise.

The second approach has taken a more empirically grounded approach in measuring the development of issue expertise. Specifically, it looks at the number of opinions authored in a given issue area, with the assumption that as one begins to author a substantial number of opinions in a given issue area then expertise begins to follow. Maltzman, Spriggs, and Wahlbeck use this measure and find that opinion-writing assignments are influenced by it.[26] Black, Johnson, and Wedeking use this measure to show that issue experts are more likely to interrupt their colleagues at oral argument, as well as the fact that Justices Blackmun and Powell were more likely to take notes at oral arguments when a Justice with more expertise spoke.[27] While this measure has a number of desirable properties (e.g., it is dynamic in that it changes over time and it varies across issue areas), we note that it is also limited because it says nothing about the quality of the opinions or the actual content contained in the opinions.

The third approach incorporates factors from a Justice's social background. Specifically, it is theorized that Justices with certain previous work or life experiences may provide them with tendencies

---

22.   John M. Scheb II & Lee W. Ailshie, *Justice Sandra Day O'Connor and the "Freshman Effect,"* 69 JUDICATURE 9, 9-12 (1985).

23.   Thea F. Rubin & Albert P. Melone, *Justice Antonin Scalia: A First Year Freshman Effect?*, 72 JUDICATURE 98, 98-102 (1988).

24.   Albert P. Melone, *Revisiting the Freshman Effect Hypothesis: The First Two Terms of Justice Anthony Kennedy*, 74 JUDICATURE 6, 6-13 (1990).

25.   Hagle, *supra* note 19, at 1147-53.

26.   FORREST MALTZMAN, JAMES F. SPRIGGS II & PAUL J. WAHLBECK, CRAFTING LAW ON THE SUPREME COURT: THE COLLEGIAL GAME 38 (2000).

27.   RYAN C. BLACK, TIMOTHY R. JOHNSON & JUSTIN WEDEKING, ORAL ARGUMENTS AND COALITION FORMATION ON THE U.S. SUPREME COURT: A DELIBERATE DIALOGUE 14-15 (2012).

to decide cases a certain way (e.g., former prosecutors tend to make more conservative decisions). A commonly cited example is Justice Harry Blackmun's authorship of the landmark abortion case *Roe v. Wade*.[28] Greenhouse argues that Blackmun drew upon his experience as a lawyer for the Mayo Clinic to help him draft the majority opinion, which is known for conceptualizing the case through the eyes of doctors, and not through the eyes of women's rights per se. While these factors are important, we note that social background factors are unlikely to be a dynamic factor in developing expertise that is able to capture the changes during the course of a Justice's tenure on the Court.

Assessing these three approaches together, we think a common strand that runs through all of them is the fact that they equate experience with expertise. While equating experience with expertise may work as a rough proxy, we think expertise is more than simply acquiring plain "experience" with the passage of time. Expertise is the ability to show mastery of a domain that novices do not possess, and that this mastery should provide benefits in the form of more skilled opinions and better legal policy. Moreover, in an environment that is rapidly changing, expertise needs to be dynamic. As a result, we think the literature is lacking a robust and dynamic measure of expertise. To search for alternative approaches, we shift to research in psychology and its long lineage of examining domain expertise.

III. EXPERTISE IN PSYCHOLOGY

Simon and Chase's foundational theory is perhaps the most important and robust theory on expertise.[29] It argues that the differentiation between experts and novices depends upon the amount of knowledge one possesses and the way that knowledge is organized.[30] Simon and Chase showed that chess masters had vastly superior recall compared to novices when they were briefly presented regular game positions.[31] It was theorized that the masters had acquired large amounts of knowledge of chess positions and, unlike novices, had organized or "chunked" the positions of individual pieces as single, meaningful configurations.[32] It was this

---

28. *See* 410 U.S. 113 (1973).
29. Chase & Simon, *supra* note 10, at 55-81.
30. *Id.* at 80-81.
31. *Id.* at 77.
32. *Id.* at 80.

organization that gives experts, but not novices, the ability to recall chess configurations as meaningful.[33]

Despite this important finding, many researchers instead focused on a different aspect of the Simon and Chase study. In addition to their finding that experts appear to chunk domain knowledge, Simon and Chase reported that chess champions had at least ten years of domain related experience before winning international tournaments.[34] This finding garnered much attention and became somewhat of a gold standard (along with reputation among peers and amount of education) by which to identify an expert.[35] However, the belief that a sufficient amount (e.g., ten years) of experience or practice leads to expertise drew criticism by scholars who showed only a weak to moderate relationship between performance level and the amount of practice and experience.[36] With evidence that experience does not necessarily mean superior performance or skill within a domain, scientists returned to consider the relationship between knowledge, structure, and expertise.[37]

Studies looking at the structural component of expertise had more luck in linking the structure of knowledge to superior skill in a number of areas.[38] More specifically, research established unique differences in the way that experts and novices organize knowledge. For example, expert knowledge representations of domain related information are superior and well defined compared to those of novices.[39]   In   particular,   experts   center   their   knowledge

---

33.   *Id.* at 80-81.

34.   *Id.* at 56.

35.   Michael I. Posner, *Introduction* to MICHELENE T.H. CHI, ROBERT GLASER & MARSHALL J. FARR, THE NATURE OF EXPERTISE, at xxix, xxxi-xxxii (1988). *See generally* THE PSYCHOLOGY OF EXPERTISE: COGNITIVE RESEARCH AND EMPIRICAL AI (Robert R. Hoffman ed., 1992).

36.   *See* Ericsson, Krampe & Tesch-Römer, *supra* note 9, at 366.

37.   To further elaborate, ten years of domain related experience may not make one an expert, but most research indicates that at least ten years of practice or experience is necessary for expertise. In other words, ten years of experience is necessary, but not sufficient, in the acquisition of expertise. *See id.* at 365-68.

38.   *See, e.g.*, Janice D. Gobert, *Expertise in the Comprehension of Architectural Plans: Knowledge Acquisition and Inference-Making*, *in* VISUAL & SPATIAL REASONING DESIGN 184, 200-05 (John S. Gero & B. Tversky eds., 1999); Egan & Schwartz, *supra* note 10, at 152-58; Ericsson & Polson, *supra* note 10; Katherine B. McKeithen et al., *Knowledge Organization and Skill Differences in Computer Programmers*, 13 COGNITIVE PSYCHOL. 307, 307-09 (1981).

39.   *See, e.g.*, Bat-Sheva Eylon & F. Reif, *Effects of Knowledge Organization on Task Performance*, 1 COGNITION & INSTRUCTION 5, 8, 10 (1984); Jill

representations on deep relationships between key domain concepts, while novices are centered on superficial connections between non-essential domain concepts.[40] In addition, network models of knowledge show that novices tend to link domain information in a more random manner compared to experts.[41]

To elaborate further, in a study of expertise in problem solving, Chi, Feltovich, and Glaser found that experts categorize problems according to "deep" features while novices categorize problems using "surface" features.[42] They asked experts and novices to sort physics word problems based on their similarity.[43] Experts sorted them according to underlying physics principles (e.g., Newton's Second Law) while novices sorted them on the basis of the surface similarity of the problem statement (e.g., both problems involved an inclined plane).[44] Furthermore, experts used their well-organized internal library of previously experienced physics problems to match the current problem with a problem solving strategy.[45] Novices, lacking such organized, extensive representations, attempted to construct solutions from the information given in the problem statement.[46]

A network study by Schvaneveldt et al. showed that the knowledge network representations of aviation concepts were most similar among expert pilots.[47] Specifically, expert pilots' networks contained only the most pertinent domain concepts, and only the

---

H. Larkin, *Processing Information for Effective Problem Solving*, 70 ENGINEERING EDUC., 285, 285-88 (1979).

40. *See* Michelene T.H. Chi, Paul J. Feltovich & Robert Glaser, *Categorization and Representation of Physics Problems by Experts and Novices*, 5 COGNITIVE SCI. 121, 121-23 (1981).

41. *See, e.g.*, Kurt Kraiger, Eduardo Salas & Janis A. Cannon-Bowers, *Measuring Knowledge Organization as a Method for Assessing Learning During Training*, 37 HUM. FACTORS 814, 810-11 (1995); William C. McGaghie et al., *Quantitative Concept Mapping in Pulmonary Physiology: Comparison of Student and Faculty Knowledge Structures*, 23 ADVANCES PHYSIOLOGY EDUC. 72, 75 (2000); Renée J. Stout, Eduardo Salas & Kurt Kraiger, *The Role of Trainee Knowledge Structures in Aviation Team Environments*, 7 INT'L J. AVIATION PSYCHOL. 235, 238-39 (1997).

42. *See* Chi, Feltovich & Glaser, *supra* note 40, at 124-25.

43. *Id.* at 130-31.

44. *Id.* at 132-33.

45. *Id.* at 132.

46. *Id.*

47. *See* Roger W. Schvaneveldt et al., *Measuring the Structure of Expertise*, 23 INT. J. MAN-MACHINE STUD. 699, 715 (1985).

most salient associations between concepts were linked.[48] Novice pilots' networks, however, varied from one another, contained non-critical domain concepts, and included seemingly random links between concepts.[49] Furthermore, the average novice network contained sixty-five links compared to the average expert network that contained thirty-nine links, indicating that novices were making extraneous connections between concepts that experts were not.[50] The knowledge networks of experts, in contrast, were sparser, representing only the most relevant domain associations.[51]

Though network studies are useful in identifying structural differences between experts and novices, typical methods of network construction require experimental conditions that are not amenable to studies of elites or others where questionnaire administration is not feasible. Fortunately, there are ways to derive networks from archival data that can help us characterize properties of experts. In particular, the relatively young field of linguistic complex networks is dedicated to the extraction of networks from textual data.[52] This field has recently seen an abundance of success in using text based networks to investigate issues in natural language processing, including the identification of literary movements,[53] text summarization,[54] and the quality of machine translation.[55] While relatively little research has examined the properties of networks derived from text to study expertise, there are two studies in the linguistic complex networks literature that are directly relevant, in that they correlate complex network properties to text quality. In a study by Antiqueira, complex networks were derived from the text of

---

48.   *See id.* at 711-13.

49.   *See id.*

50.   *Id.*

51.   *See id.* More recently, however, Schuelke et al. found that the number of network links was not meaningfully correlated with level of performance skill. *See* Matthew J. Schuelke et al., *Relating Indices of Knowledge Structure Coherence and Accuracy to Skill-Based Performance: Is There Utility in Using a Combination of Indices?*, 94 J. APPLIED PSYCHOL. 1076, 1082 (2009).

52.   *E.g.*, Jin Cong & Haitao Liu, *Approaching Human Language with Complex Networks*, 11 PHYSICS LIFE REVS. 598, 599 (2014).

53.   *See* Diego Raphael Amancio, Osvaldo N. Oliveira Jr. & Luciano da Fontoura Costa, *Identification of Literary Movements Using Complex Networks to Represent Texts*, NEW J. PHYSICS Apr. 2012, at 1, 2.

54.   *See* Diego R. Amancio et al., *Complex Networks Analysis of Language Complexity*, EUROPHYSICS LETTERS, Dec. 2012, at 1, 1; Lucas Antiqueira et al., *A Complex Network Approach to Text Summarization*, 179 INFO. SCI. 584, 585 (2009).

55.   *See* D.R. Amancio et al., *Using Metrics from Complex Networks to Evaluate Machine Translation*, 390 PHYSICA A 131, 137 (2011).

forty high school student essays written on the same topic.[56] In addition, five human judges independently scored the essays on the basis of text quality.[57] Antiqueira found that both average node strength and average clustering coefficient correlated negatively with text quality.[58] In a similar study, Ke derived complex networks using a set of 500 short-answer responses from a Chinese achievement exam.[59] The essays were additionally scored for quality by human raters.[60] Like Antiqueira, Ke found that average clustering coefficient and average node strength decrease with increasing text quality, and additionally found that the number of network nodes decreases with increasing text quality.[61] Findings from these two studies provide some insight into what network indicators will be helpful in identifying latent factors of judicial expertise.

A common theme of early studies of domain expertise is that the domain space is stable (e.g., the domain of chess expertise does not change because the rules are fixed). In cases where this holds, traditional measures of expertise may be suitable. However, many domains are not stable and require a different approach to the study of expertise. In fact, Johansson argues that many "principles" of what defines an expert go "out the window" in ever-changing, less stable domains.[62] A recent study by Macnamara et al. supports the idea that the study of expertise may be domain dependent.[63] In a meta-analysis of eighty-eight studies on deliberate-practice, Macnamara and colleagues found that practice accounted for only 12% of the difference in performance in a number of domains.[64] This means that some of the defining principles of expertise (i.e., Ericsson's 1993 theory of deliberate practice) are not applicable or are not as critical across domains.[65] This suggests that network

56. L. Antiqueira et al., *Strong Correlations Between Text Quality and Complex Networks Features*, 373 PHYSICA A 811, 812 (2007).

57. *Id.*

58. *Id.* at 817.

59. Xiaohua Ke et al., *Complex Dynamics of Text Analysis*, 415 PHYSICA A 307, 308 (2014).

60. *Id.*

61. *See id.* at 312.

62. *See* FRANS JOHANSSON, THE CLICK MOMENT: SEIZING OPPORTUNITY IN AN UNPREDICTABLE WORLD 22-32 (2012).

63. Brooke N. Macnamara, David Z. Hambrick & Frederick L. Oswald, *Deliberate Practice and Performance in Music, Games, Sports, Education, and Professions: A Meta-Analysis*, 25 PSYCHOL. SCI. 1608, 1615 (2014).

64. *Id.* at 1610, 1612.

65. *Id.* at 1615.

structure alone may not be enough to characterize expertise in a dynamic domain. To make sure we account for the major components of expertise in the legal domain, it is necessary to consider how Justices adapt to the changing environment, so we need to evaluate how network structure changes over time. In addition, we also need to include a component to measure the amount of information a Justice displays, in order to account for the information asymmetry between experts and novices.

The work cited above suggests it is possible to distinguish between experts and novices based on network properties. Furthermore, this approach offers advantages over traditional measures of expertise in that it is applicable to a broad range of subjects, assuming there is corresponding textual data (e.g., speeches, opinions, and communications). What is needed now is an approach that takes our network approach to the study of expertise and adds a component that accounts for the information advantage that experts have over novices.

IV. A NEW MEASURE OF EXPERTISE

Measuring judicial expertise presents several challenges. First, the measure must be dynamic in that it accounts for the changing environment and also any changes in how a Justice adapts (e.g., writing shorter or longer opinions). Second, a new measure must be able to be applied to elites and those that are unwilling to answer questionnaires. Third, to understand how expertise has functioned in other time periods, it must be applicable to other eras, even when the subjects of interest are dead. To meet these challenges, we apply networks to textual data. Specifically, we derive co-occurrence networks from Supreme Court opinions, and use features of these networks to characterize legal expertise. We do this because of previous research showing that parameters of linguistic networks are potential indicators of differences in content and style of text.[66]

Our first step in constructing co-occurrence networks from Supreme Court opinions involves pre-processing the text. More specifically, we remove *stop words*, which are words that have little semantic value.[67] Next, the remaining words are modeled as nodes,

---

66.   *See generally* Jin Cong & Haitao Liu, *Approaching Human Language with Complex Networks*, 11 PHYSICS LIFE REVS. 598 (2014).

67.   Most of the words considered stop words are articles and prepositions (e.g., the, it, or, of) and are removed because they function in the text namely to connect other content words.

and nodes are connected with a directed, weighted link if they occur in a window of two words within the text (e.g., they are adjacent in the text). Link weights are simply the frequency of the association between the two words throughout the unit of text, assuming word order is preserved. Thus, if two words co-occur three times throughout a text, their weight would be three. Link directions are defined by natural reading order, where the word on the left is the source node, and the substantive word immediately to the right is the target node.

To illustrate how we construct a co-occurrence network, Figure 1 shows an example of the corresponding co-occurrence network extracted from this famous legal quote: "[T]he dichotomy between personal liberties and property rights is a false one. Property does not have rights. People have rights."[68] All links in Figure 1 are weighted with 1 because the words co-occur only once, with the exception of the link between "property and rights" that co-occurs twice in the text. The unit of analysis was a single majority opinion, and so each co-occurrence network represented the text of a single majority Supreme Court opinion. Thus, while the example is only a mere three sentences, the same principles are applied when extending the method to longer documents.[69]

---

68. Lynch v. Household Fin. Corp., 405 U.S. 538, 552 (1972).

69. We focused only on majority opinions because they are generally known to follow a template that is similar across Justices (e.g., facts of the case, followed by the reasoning). Dissents and concurring opinions, however, do not follow a consistent format or style. Only examining majority opinions ensures that differences in the written structure are more likely attributable to the author's expertise and less to the type of opinion. Another concern is that some might argue the clerks do a large part of the opinion writing or at least draft major portions of an opinion. While there are some anecdotes, *see* BOB WOODWARD & SCOTT ARMSTRONG, THE BRETHREN: INSIDE THE SUPREME COURT (1979), and survey evidence, *see* ARTEMUS WARD & DAVID L. WEIDEN, SORCERERS' APPRENTICES: 100 YEARS OF LAW CLERKS AT THE UNITED STATES SUPREME COURT (2006), that some Justices rely more on their clerks than others, statistical evidence trying to capture the linguistic "fingerprints" of clerks is less supportive. *See* Paul J. Wahlbeck, James F. Spriggs II & Lee Sigelman, *GhostWriters on the Court?: A Stylistic Analysis of U.S. Supreme Court Opinion Drafts*, 30 AM. POL. RES. 166, 182-83 (2002); IAN SULAM, EDITOR IN CHIEF: OPINION AUTHORSHIP AND CLERK INFLUENCE ON THE SUPREME COURT 19 (2014), http://icsulam.github.io/pdf/EditorInChief.pdf [https://perma.cc/Z6GB-KVXS]. Regardless, it is important to note that at a minimum, all Justices read, approve, and have final say over the opinions that they author.
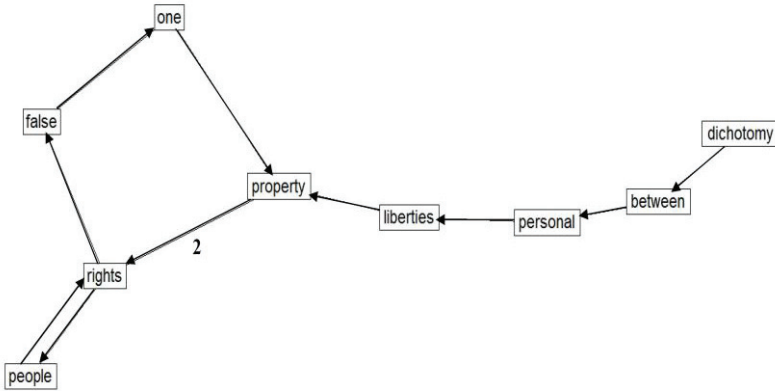
Figure 1. Example co-occurrence network derived from the text: "[T]he dichotomy between personal liberties and property rights is a false one. Property does not have rights. People have rights."[70] The link weights between the terms are all one, except for the weight of 2 between "property" and "rights", which co-occur twice in the text.

After we constructed a network of N nodes from an opinion, we calculated the relevant network measures from its corresponding directed and weighted adjacency matrix W. The $N \times N$ adjacency matrix was obtained by starting with all zero elements and setting $W(i,j) = W(i,j) + 1$ whenever there was an association of node i to node j.

V. TEXT-BASED INDICATORS OF EXPERTISE

The goal is to gather as many text-based indicators of the two different factors that previous research suggests are indicative of expertise: the amount of knowledge and its structure. To construct a measure of expertise, we use text-derived measures of expertise, with twelve measures from network science, and five measures from information science that, we argue, capture components of expertise found in human studies. To that end, we turn to the respective literatures to identify the relevant indicators.

---

70.   *Lynch*, 405 U.S. at 552.

A. Network Indicators[71]

The first eight indicators from network science are standard measures that are used to study a variety of natural language processing tasks, including document classification,[72] text summarization,[73] authorship identification,[74] and quality analysis.[75] In addition, these measures have been applied to studies ranging from cell biology to business.[76] These eight network measures are: number of nodes, number of links, clustering coefficient, node strength, average shortest path length, efficiency, diameter, and radius. We discuss each in turn.

1. *Number of Nodes*

The number of nodes, N provides an approximation for the number of unique concepts in a network. Recall that chunking allows experts to store multiple superficial features into a single deep feature.[77] In the linguistics domain, Ke found that text quality was lower in networks with large N.[78] We expect to see similar results in our work; namely, that as judicial expertise increases, the number of nodes will decrease in the representative co-occurrence networks.

---

71.   In non-trivial cases of computation, we provide a reference for the equation used.

72.   *See* Rada Mihalcea & Dragomir Radev, Graph-Based Natural Language Processing and Information Retrieval 98 (2011); *see also* Sabina Šišović, Sanda Martinčić-Ipšić & Ana Meštrović, *Comparison of the Language Networks from Literature and Blogs*, *in* 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) 1603 (Petar Biljanovic et al. eds., 2014).

73.   *See* Antiqueira et al., *supra* note 54, at 584.

74.   *See* Efstathios Stamatatos, *A Survey of Modern Authorship Attribution Methods*, 60 J. Am. Soc'y for Info. Sci. & Tech. 538, 538 (2009).

75.   *See* Antiqueira et al., *supra* note 56, at 812; *see also* Ke et al., *supra* note 59, at 308.

76.   *See* S. Boccaletti et al., *Complex Networks: Structure and Dynamics*, 424 Physics Rep. 175, 260-71 (2006); *see also* Steven H. Strogatz, *Exploring Complex Networks*, 410 Nature 268, 268 (2001).

77.   *See* Chase & Simon, *supra* note 10, at 56; *see also* Chi, Feltovich & Glasser, *supra* note 40, at 122.

78.   Ke et al., *supra* note 59, at 310.

2. *Number of Links*

The number of links is the number of node-to-node associations that exist in the network. Cognitive psychology studies suggest that knowledge networks derived from experts maintain only the most salient links between concepts, whereas knowledge networks derived from novices tend to have more superficial, random associations between concepts.[79] Similarly, Ke found co-occurrence networks of high quality text had fewer links than co-occurrence networks from low quality text.[80] Both findings suggest a sparseness to co-occurrence networks representing expert knowledge, and thus predict that expert co-occurrence networks from expertly written text should be leaner compared to their novice counterparts.

3. *Clustering Coefficient*

The clustering coefficient quantifies the probability that a given node's neighbors are neighbors themselves.[81] Let k_i be the number of neighbors of node i, and w_{ij} be the weight of the link from node i to node j, then the weighted clustering coefficient of node i is given by

$$C_i = \frac{2}{k_i(k_i - 1)} \sum_{(j,k)} (\widehat{w}_{ij}\widehat{w}_{jk}\widehat{w}_{ki})^{1/3}$$

where $\widehat{w}_{ij} = w_{ij}/max_{ij}w_{ij}$.

The average clustering coefficient of the network is used in this paper, and is the arithmetic mean of all individual clustering coefficients:

$$C = \frac{1}{N} \sum_{i}^{N} C_i$$

---

79.   *See* Schvaneveldt et al., *supra* note 47, at 701-02.
80.   *See* Ke et al., *supra* note 59, at 311.
81.   Giorgio Fagiolo, *Clustering in Complex Directed Networks*, PHYSICAL REV. E, Aug. 2007, at 1, 2.

With respect to text, the clustering coefficient indicates the interconnectedness of concepts, which arguably reflects the degree of interconnection between concepts in the author's knowledge representation. Linguistic studies show that co-occurrence networks derived from poor quality text have high average clustering coefficients and those derived from high quality text have low average clustering coefficients.[82] Experimental studies of expertise demonstrate a similar finding, in that experts have fewer, but more salient links between concepts where novices have more links, and such links connect random concepts.[83]

4. *Node Strength*

A single node has a strength equal to the sum of its weights, and the average node strength is the arithmetic mean of all individual node strengths.

$$str = \frac{1}{N} \sum_{i=1}^{N} w(j, i)$$

Node strength is another measure of interconnectedness, and so it is not surprising that node strength and clustering coefficient reflects the same patterns with respect to expertise. Indeed, Antiqueira and Ke found that text quality decreases as average node strength increases.[84] Šišović used complex network features to discriminate between novels and blogs, and showed that average node strength is less for novels than for blogs.[85] These results become intuitive if we assume blogs are samples of novice writing and novels are samples of expert writing.

5. *Shortest Path Length*

A *path* is the particular route taken to go between any two nodes in a network. For any two nodes in a network, there exists a path that requires the minimum number of links needed to travel

---

82. *See* Antiqueira et al., *supra* note 56, at 815-19; *see also* Ke et al., *supra* note 59, at 311.
83. *See* Schvaneveldt et al., *supra* note 47, at 701-02.
84. *See* Antiqueira et al., *supra* note 56, at 815-19; *see also* Ke et al., *supra* note 59, at 311.
85. *See* Šišović, Martinčić-Ipšić & Meštrović, *supra* note 72, at 1608.

between the two. This path is known as the shortest path length between the two nodes. Let d_{ij} give the length of the shortest path (minimal length) that connects node *i* to node *j*. The average shortest path length, then, for a network with *N* nodes is given by

$$L = \frac{1}{N(N-1)} \sum_{\substack{i \neq j}}^{N} d_{ij}$$

In complex network studies of language, the average shortest path length has been indicated as measuring the extent to which the author is able to establish long sequences of connections among different concepts.[86] A small average shortest path length is indicative of poor writing, in that it reflects an author's difficulty in making associations between semantically distant concepts.[87] In contrast, experts have more extensive domain knowledge compared to novices,[88] indicating a larger capability than novices in associating semantically distant concepts. Likewise, in the linguistics literature,[89] Antiqueira and Margan both found that text quality and meaningfulness of text, respectively, were positively correlated with the size of the average shortest path length of corresponding co-occurrence networks.

6. *Efficiency*

The average or global efficiency of a network of *N* nodes is proportional to the sum of the inverse shortest paths.

$$E = \frac{1}{N(N-1)} \sum_{\substack{i \neq j}}^{N} \frac{1}{d_{ij}}$$

Because average shortest path and efficiency are inversely related, we expect efficiency to increase as text quality and meaningfulness

---

86. *See* Antiqueira et al., *supra* note 56, at 818.
87. *Id.*
88. *See, e.g.*, Nigel R. Shadbolt & Paul R. Smart, *Knowledge Elicitation: Methods, Tools and Techniques*, *in* EVALUATION OF HUMAN WORK 163 (John R. Wilson & Sarah Sharples eds., 4th ed. 2015).
89. *See* Antiqueira et al., *supra* note 56, at 818; Domagoj Margan, Sanda Martinčić-Ipšić & Ana Meštrović, *Network Differences Between Normal and Shuffled Texts: Case of Croatian*, *in* COMPLEX NETWORKS V 275, 282 (Springer International Publishing 2014).

decrease. In turn, this implies that experts' knowledge representations will be characterized as less efficient compared to novices' networks because they have larger representations to traverse.

7. *Diameter*

Given the set of all shortest paths between all connected nodes in a network, the diameter is the maximum of this set. Given the more extensive knowledge representations of experts compared to novices, we expect that the maximum shortest path length to be larger for networks representing expert knowledge than for networks representing novice knowledge. Margan supports this notion, finding a smaller diameter for co-occurrence networks from shuffled (word order was randomized) versus unshuffled text.[90] In other words, when text becomes less meaningful, its representative co-occurrence network will correspondingly reduce in diameter.

8. *Radius*

Where diameter is the maximum of the set of all shortest paths between connected nodes of a network, the radius is the minimum of this set. We can infer correlations between radius and expertise from the findings on average shortest path length and text quality. In particular, more expertly written opinions should exhibit a larger radius due to experts' more extensive knowledge. When comparing the average network derived from text of an expert to the average network from text of a novice, we expect that the novice network is less extensive and so has a smaller radius.

The final four network measures are standard measures that have been important for various research applications[91] and show theoretical promise to being relevant to expertise. These network metrics are betweenness centrality, community affiliation, modularity index, and node diversity.

---

90.   *See* Margan, Martinčić-Ipšić & Ana Meštrović, *supra* note 89, at 282.

91.   *See generally* ALBERT-LÁZLÓ BARABÁSI & JENNIFER FRANGOS, LINKED: THE NEW SCIENCE OF NETWORKS (2014).

9. *Betweenness Centrality*

Betweenness centrality measures the extent to which a node lies on paths between other nodes.[92] For our purposes, we calculate the average betweenness centrality over all network nodes. Networks derived from expertly written text should have lower average betweenness centrality than the networks derived from novice text because, in general, their knowledge representations are not as dense and connect important concepts directly.[93] In this way, they should not require "in between" nodes to associate two concepts. However, note that linguistic research shows inconclusive results concerning text quality and betweenness centrality.[94]

10. *Community Affiliation*

A network is said to have community structure if it can be easily grouped into sets of nodes such that each set of nodes is densely connected internally and sparsely connected externally.[95] Following Antiqueira and colleagues,[96] we assume that communities correspond to the topics conveyed by the text. While studies using co-occurrence networks have not shown a conclusive relationship between text properties and community affiliation measures, studies of text quality show that the coherence or fluidity of the text as judged by human raters increases with the domain knowledge of the writer.[97] Furthermore, expert/novice studies suggest that knowledge of experts is more coherent compared to novices.[98] That is, experts demonstrate more consistency in the associations they make between concepts stored in memory.[99] In this way we expect community affiliation to decrease with increasing expertise.

---

92. Ulrik Brandes, *A Faster Algorithm for Betweenness Centrality*, 25 J. MATHEMATICAL SOC. 163, 167 (2001).

93. *See* Schvaneveldt et al., *supra* note 47, at 713.

94. Amancio et al., *supra* note 54, at 6; *see* Šišović, Martinčić-Ipšić & Meštrović, *supra* note 72, at 1606-08.

95. E.A. Leicht & M.E.J. Newman, *Community Structure in Directed Networks*, PHYSICAL REV. LETTERS, Mar. 2008, at 1, 1.

96. *See* Antiqueira et al., *supra* note 54, at 584.

97. *See* Šišović, Martinčić-Ipšić & Meštrović, *supra* note 72, at 1603; Ronald T. Kellogg, *Long-Term Working Memory in Text Production*, 29 MEMORY & COGNITION 43 (2001).

98. Stout, Salas & Kraiger, *supra* note 41, at 246.

99. Timothy Goldsmith & Kurt Kraiger, *Structural Knowledge Assessment and Training Evaluation*, *in* IMPROVING TRAINING EFFECTIVENESS IN WORK

11. *Modularity Index*

Presupposing an optimal community structure of a network, the modularity index quantifies the degree to which the network may be subdivided into these clearly delineated groups.[100] A high modularity index means that the network is easily subdivided, and these divisions are relatively obvious. Because the modularity index depends upon the community affiliation measure, we use the same logic in predicting the modularity index's relation to expertise—expertly written opinions should be less segmented, making their divisiveness low, yielding a low modularity index.

12. *Node Diversity*

Node diversity is another measurement of community distinction, and the average node diversity is simply the arithmetic mean over all node diversity values in the network.[101] We draw upon our arguments from the other community distinction measures (i.e., community affiliation and modularity index) to suggest that expert knowledge that is highly fluid and coherent will be represented in highly fluid and coherent text. This means that co-occurrence networks derived from expert writings should have low node diversity, while co-occurrence networks derived from novice writings will have comparatively higher node diversity.

B. Cognitive Science Indicators

A large body of research suggests that the amount of information is increasing and making the external world more complex. Furthermore, experts have greater amounts of knowledge and information than novices, so we need to account for this disparity in our measure. Thus, we next describe four cognitive science indicators taken for this purpose: review internal entropy (RI-Ent), average unigram information (AUI), average conditional

---

ORGANIZATIONS 73-74 (J. Kevin Ford ed., 1997); Stout, Salas & Kraiger, *supra* note 41, at 246.

    100.   *See generally* Leicht & Newman, *supra* note 95, at 11; Jörg Reichardt & Stefan Bornholdt, *Statistical Mechanics of Community Detection*, PHYSICAL REV. E, July 2006, at 1.

    101.   Mikail Rubinov & Olaf Sporns, *Weight-Conserving Characterization of Complex Functional Brain Networks*, 56 NEUROIMAGE 2068, 2072 (2011); C.E. Shannon, *A Mathematical Theory of Communication*, 27 BELL SYS. TECH. J. 379 (1948).

information (ACI), and conditional information variability (CIV). These four indicators were adapted from Vinson and Dale's work, which looked at informational content contained in Yelp user reviews.[102] Each of the four functions presents a slightly different take on the amount of information contained in a unit of text and may, in turn, reveal unique insights into expertise. We also include the average number of words in an opinion each term for our fifth information measure.

1. *Review Internal Entropy (RI-Ent)*

The user "review" studied by Vinson and Dale is analogous to a single opinion for our purposes. RI-Ent measures the amount of information contained in a single opinion in terms of the novelty of the language use within the opinion.[103] In other words, if the text is highly repetitive (e.g., The cat in the hat. The cat in the hat sat) it will have a lower RI-Ent score, but if the text contains more lexical richness (e.g., The cat in the hat. A dog came over to play), it will have a higher RI-Ent score. The following equation mathematically represents this lexical richness:

$$RI - Ent_j = -\sum_{i=1}^{N} p(w_i|O_j) \log_2 p(w_i|O_j)$$

Here, RI-Ent$_j$ yields the information contained within the jth opinion, containing N words, as a function of the probability of the ith word occurring within that opinion (treated as the conditional probability that given opinion j, word i will be present). The RI-Ent score quantifies the uniqueness of the language in an opinion because it increases as the number of novel words increases.

2. *Average Unigram Information (AUI)*

Like RI-Ent, AUI is a measure of lexical richness of a single opinion.[104] However, it compares the probability of occurrence of a word in an opinion to its probability of occurrence across the entire

---

102.   David W. Vinson & Rick Dale, *Valence Weakly Constrains the Information Density of Messages*, *in* PROCEEDINGS OF THE 36TH ANNUAL CONFERENCE OF THE COGNITIVE SCI. SOC'Y 1682 (2014).
103.   *See id.* at 1683-84.
104.   *Id.* at 1684.

corpus of text used in this study.[105] The information encoded in a word can be given as the negative log of the probability of its occurrence (the less probable a word, the more informative), for any given opinion j, the following equation defines its average unigram information by averaging over the probability of occurrence of the ith word in the opinion:

$$AUI_j = -\frac{1}{N}\sum_{i=1}^{N} \log_2 p(w_i)$$

In this way, AUI gives a direct measure of the relative informational content of an opinion across all opinions and may vary depending on the expertise of the Justice.

3. *Average Conditional Information (ACI)*

ACI measures how informative a word is given its local context.[106] In other words, it measures how likely it is that a given word appears in text, considering the previous word in text.[107] In this way, ACI acts as a measure of the uniqueness of bigrams in text, where uniqueness is measured relative to the distribution of bigram (two-word) frequencies in the corpus. Mathematically, the conditional information in opinion *j* is the average negative log of the probability of a word's occurrence given the previous word, as defined in the following equation:

$$ACI_j = -\frac{1}{N-1}\sum_{i=1}^{N} \log_2 p(w_i|w_{i-1})$$

4. *Conditional Information Variability (CIV)*

Unlike RI-Ent, AUI, and ACI, which indicate the average amount of information in an opinion, CIV quantifies the degree to which the language of an opinion varies in its informational content by calculating the variability in the ACI measure.[108] It is expressed as the standard deviation of the set of conditional information scores for each word of the jth opinion:

$$CIV_j = \sigma(CI_j)$$

105. *Id.*
106. *Id.*
107. *See id.*
108. *See id.* at 1683-85.

5. *Average Number of Words*

In addition to RIE-Ent, AUI, ACI, and CIV, we look to Posner who says experts should be able to "say more with less" and take as our fifth cognitive indicator (and seventeenth and final indicator of expertise), the average number of words per opinion each term.[109]

Using these seventeen indicators, because these measures are continuous (and not categorical), we employ factor analysis to detect underlying latent factors of expertise.

VI. METHODS

To best study judicial expertise and its dynamic nature, we chose to study Supreme Court Justices because of the importance of the Supreme Court as an institution in American government. Furthermore, because the Supreme Court only has nine Justices at any given time and because the memberships of the nine Justices do not overlap uniformly, for our study we chose to examine any Justice who has been on the Supreme Court for a minimum of ten terms. Because a study of all issue areas is not feasible, we chose to focus our attention on the majority opinions from two issue domains identified by the Supreme Court Database: criminal procedure and economic activity.[110] Those two issue domains comprise a very large number of cases and provide the best assurances that each Justice typically wrote at least one opinion in that area each term. Ideally, we would like to have a Justice write several opinions on each issue area each term, but the practicalities of nonrandom case selection at the Supreme Court coupled with the rules and workload concerns regarding opinion writing assignments limit the number of opinions written by each Justice within each issue area.[111]

---

109.   *See* POSNER, *supra* note 1, at 297-301.

110.   *See Analysis Specifications*, THE SUP. CT. DATABASE, http://supremecourtdatabase.org/analysis.php (last visited Apr. 5, 2016) [https://perma.cc/X22Z-AARK]. Criminal procedure involves issues such as: habeas corpus, search and seizure, *Miranda* warnings, right to counsel, cruel and unusual punishment, and other related issues (includes all issues in the Database that receive a code from 10000 to 10600). Economic activity issues include: antitrust, mergers, bankruptcy, liability, securities regulation, patents and copyrights, and other related issues (includes all issues in the Database that receive a code from 80000 to 80350).

111.   For example, the Chief Justice and most senior Associate Justices tend to assign the bulk of who writes the majority opinion. While there can be some strategy involved in opinion assignment, workload considerations also play a significant role.

We then construct networks from the written opinions following our method outlined above and then use that data to create measures of expertise for each term a Justice is on the Court. So if a Justice wrote three majority opinions in one term on criminal procedure, the expertise indicators would be an average taken from those three opinions.

After we have the indicators, the next step performs factor analysis on the seventeen indicators from the criminal procedure area to identify the two factors that characterize expertise.[112] Table 1 contains the results of the two-factor solution for the criminal procedure domain (eigenvalues were 8.55 and 2.32, respectively), with an orthogonal rotation.[113] A commonly accepted rule-of-thumb is that any factor that loads at greater than .3 suggests that an indicator loads onto that factor. With this in mind, the first two columns of loadings in Table 1 show that both factors are well defined in the sense that each factor has at least four indicators loading on each factor separately, though some indicators do load onto both factors. While this may not be ideal, it is not concerning because each factor has enough individual items to be sufficiently defined. Furthermore, the uniqueness values in the far right column are all generally low, with the possible exception of radius, modularity index, ACI, and CIV, though their uniqueness values are not concerning because each of their loadings are at least .34 or greater.

---

112.   We performed iterated principal axis factoring, keeping all factors with eigenvalues greater than 1. This is known as Kaiser's criterion and is a widely accepted rule-of-thumb for selecting the appropriate number of factors. Henry F. Kaiser, *The Application of Electronic Computers to Factor Analysis*, 20 EDUC. & PSYCHOL. MEASUREMENT 141, 146 (1960).

113.   The orthogonal rotation assumes that the two factors are not correlated with each other. Substantively this means that the amount of knowledge is unrelated to how a person stores or structures their knowledge. We think this is a reasonable assumption under the idea that it is possible to have large amounts of knowledge but stored poorly or idiosyncratically. However, alternative analyses that allow for the two factors to be correlated return very similar results. On a related yet separate issue, the next highest eigenvalue was .694, which suggests a two-factor solution was the correct solution. Finally, to ensure that comparing Justices from different time periods did not produce a confounding effect, we also divided the data at the mid-point to construct two datasets, one for pre-1981 and one for post-1981. Factor analysis on these two time periods each revealed a two-factor solution with factor scores that correlate very highly with scores from the full time period.

Table 1
FACTOR ANALYSIS OF 17 INDICATORS OF EXPERTISE

	Amount of Knowledge	Structure	Uniqueness
Clustering Coefficient	.162	-.639	.566
Node Strength	.603	-.735	.097
Betweenness Centrality	-.792	.357	.245
Node Diversity	-.011	-.647	.582
Community Affiliation	.766	-.125	.398
Modularity Index	-.372	.297	.773
Average Shortest Path	-.472	.838	.076
Efficiency	.400	-.877	.071
Number of Links	.886	-.347	.094
Radius	.113	.405	.823
Diameter	-.099	.660	.555
Number of Nodes	.957	-.233	.031
Number of Words	.853	-.387	.124
RI-Ent	.917	-.086	.151
AUI	-.518	.797	.098
ACI	-.126	.529	.705
CIV	.363	.348	.747

Note: Table contains the rotated factor loadings after varimax rotation.

The first factor appears to be capturing what we label the "amount of knowledge" factor because the RI-Ent measure loads on it as well as the number of nodes, number of links, and number of words (recall RI-Ent captures novel information or distinct word usage within an opinion itself). This seems intuitive in that as

opinions become longer, these indicators would capture that feature. Moreover, because all of those indicators load positively onto the factor, this suggests that high values for the amount of knowledge actually correspond to low expertise (i.e., for RI-Ent, more novel information is less expert-like in that it emphasizes a focus on new or novel information, not established knowledge, which is what experts in the past were shown to focus on).

We label the second factor as "structure" because many of the indicators that tap into interconnectedness of concepts load highly on it. Specifically, these include the clustering coefficient, node diversity, and average shortest path length (and its derivatives—diameter, radius, and efficiency). ACI also loads highly on this structure factor, and it is believed to indicate the novelty of a linguistic phrase (two nodes connected by a link), which may reflect how often an opinion repeats the same connected ideas.[114]

Next, we use the factor analysis to estimate factor scores to represent the two latent constructs of expertise. This gives us a variable that represents the amount of knowledge in a domain and a variable that represents the structure of knowledge in a domain. For ease of interpretation, we multiplied the "amount" variable by -1 to ensure a more intuitive interpretation (e.g., high values would indicate more expertise).[115] Figures 2 and 3 contain a dot that is an estimate for the degree of each Justice's expertise in each year for cases in the criminal procedure issue area. The line is a lowess smoother designed to fit a weighted moving average, which should help in trying to identify any trends. Figure 2 contains the amount of knowledge factor, and Figure 3 contains what we argue represents the degree of expertise in how knowledge is structured.

---

114. The same factor analysis was performed on majority opinions from the economic issue area. The results are very similar, revealing a two-factor solution with almost all indicators loading onto the same factors. Only three of the indicators differed in their loadings: The clustering coefficient did not load highly on either factor, and the AUI and ACI indicators switched factors. Some minor deviations are to be expected given it is a different issue area and also due to the tendency for textual data to be noisy. Importantly, on the whole, the factor analyses revealed the same result.

115. Recall, this was done based on how the indicator variables loaded onto each factor in Table 1. In other words, the loadings for the "amount" factor variable initially suggested that "low values" indicate expertise. Thus, we reverse coded it to make it more intuitive to interpret.
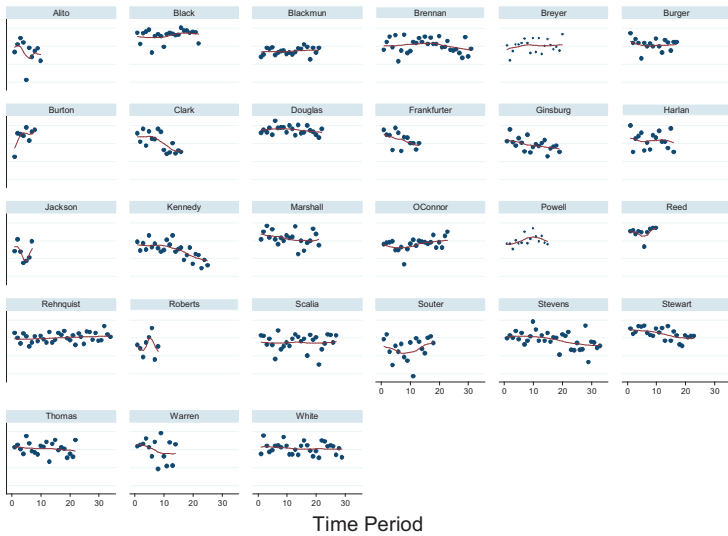
Figure 2. Amount of Knowledge Expertise Over Time, by Justice. Each dot represents the estimated level of expertise in a given year for that Justice. The line is a lowess smoother.

Figure 3. Structure of Knowledge Expertise Over Time, by Justice. Each dot represents the estimated level of expertise in a given year for that Justice. The line is a lowess smoother.
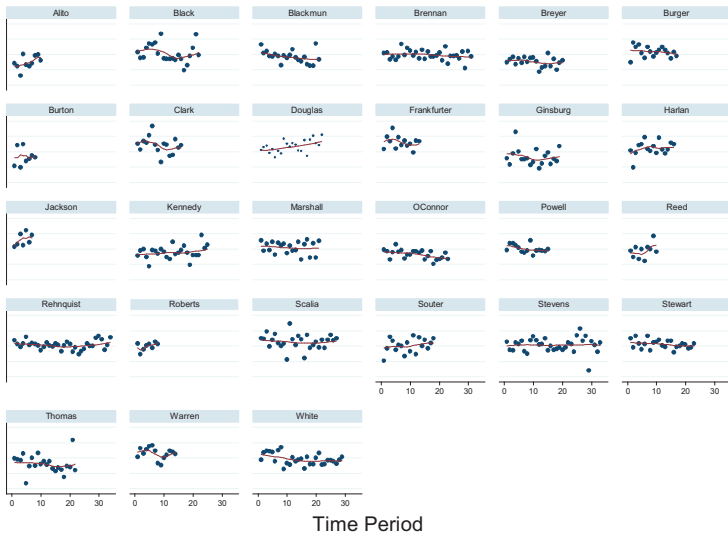
The figures reveal several important and somewhat surprising findings. First, expertise for some Justices increases over time, for some it does not change, but for others it decreases. Specifically, with respect to Figure 2, which displays the factor of expertise representing the amount of knowledge, Justices Black, Breyer,

Burton, O'Connor, Powell, Reed, and Souter all display some increased level of expertise during their tenure on the Court. In contrast, Justices Alito, Brennan, Clark, Douglas, Frankfurter, Ginsburg, Jackson, Kennedy, Stevens, Stewart, Thomas, and perhaps Warren show a decreased level of expertise. Still other Justices show no real trend in either direction: Blackmun, Rehnquist, and Scalia. Only a few Justices display a large amount of noise, such as Roberts, Harlan, and possibly Warren.[116]

We see similar trends in Figure 3 for the structure of expertise. Justices Alito, Douglas, Harlan, Jackson, Kennedy, Reed, and Souter display signs of increasing expertise over time in the form of a more expert-like structure. In contrast, Black, Frankfurter, O'Connor, Thomas, Warren, and White show decreasing expertise. What is interesting, when looking at Figures 2 and 3 in conjunction with each other, we see that for some Justices whose decrease in expertise in one area was compensated for an increase in the other type (e.g., Alito, Douglas, Kennedy). This suggests that some Justices may cope with the growing complexity by developing and relying on one type of expertise more than (and maybe at the expense of) another type.

Perhaps the most important finding from Figures 2 and 3 is the fact that some Justices increased their expertise while others decreased their expertise. This is noteworthy because it confirms our earlier suspicions that expertise can decline, even as one gains more experience.

In the next section, we use our estimates of expertise to see if Justices with more expertise can reduce the number of concurring opinions.

VII. ARE EXPERTS ABLE TO REDUCE THE NUMBER OF CONCURRING OPINIONS?

To assess the validity of our two measures of expertise, we assess whether increased expertise correlates with the number of concurring opinions written in that case. Importantly, the Supreme Court is a rule-making body whose goal is to clarify the law for the lower courts and its other audiences. Thus, the Court plays a key role in crafting legal policy. Generally speaking, a "good" majority

---

116.   If a Justice does not have at least ten dots on their respective figure, it means that they did not author a criminal procedure majority opinion during that term (e.g., Roberts only has eight dots).

opinion will communicate its intent and meaning in a clear signal without ambiguity. However, one possible way this clarity can be muddied is through other Justices writing concurring opinions. This includes the possibility of either a regular concurrence or a special concurrence. A regular concurrence is when a Justice agrees with the rationale and disposition of a case and simply wants to say something in addition to the majority opinion.[117] A special concurrence is when a Justice agrees with the disposition but not the rationale.[118] Both types of concurring opinion have the capability to muddy the waters with respect to how the Court is perceived and whether the Court speaks with a singular voice. Thus, a majority opinion author wants to reduce the number of concurring opinions.[119] It is in this context that the expertise of a Justice plays a key role in shaping legal policy in America.

Importantly, Justices with more expertise should be able to limit the number of concurring opinions because Justices have the ability to craft and frame the opinions in an expert-like manner so as to better address counterarguments and explain the majority's rationale.[120] Also, because the Justices wait to see the first draft of the majority opinion before they formally decide if they will join the opinion, it enables us to test whether Justices with more expertise can limit or reduce the number of concurring opinions being written.[121]

To measure our dependent variable—the number of concurring opinions being written—we use data from the Supreme Court Database and simply count the number of concurring opinions written in each case. We also differentiate between regular concurrences and special concurrences. Thus, we have two dependent variables: the number of regular concurrences and the number of special concurrences. Each variable ranges from 0 to 4, and each is heavily skewed, with the modal outcome of both types of concurrence being "0."[122] Because our dependent variables are count variables that are over-dispersed (i.e., the standard deviation is larger

---

117. Paul M. Collins, Jr., *Amici Curiae and Dissensus on the U.S. Supreme Court*, 5 J. EMPIRICAL LEGAL STUD. 143, 153 (2008).
118. *Id.*
119. MALTZMAN, SPRIGGS & WAHLBECK, *supra* note 26, at 68-69.
120. *See* BLACK, JOHNSON & WEDEKING, *supra* note 27, at 54.
121. MALTZMAN, SPRIGGS & WAHLBECK, *supra* note 26, at 66-67.
122. For regular concurrences, the distribution of observations is: (0) 1,138; (1) 321; (2) 72; (3) 7; and (4) 3. For special concurrences, the distribution of observations is: (0) 1,111; (1) 337; (2) 71; (3) 18; (4) 4.

than the mean), we use maximum likelihood to fit a negative binomial regression model.[123]

The two measures of expertise—amount of knowledge and structure of knowledge—are our main variables of interest. For expertise, high values indicate more expertise. To ensure that we are accounting for other possible explanations of the number of concurring opinions, we include several more covariates. We include a measure if a precedent was altered, with the expectation that the rule of law is predicated upon upholding past precedents. Thus, Justices do not regularly stray from established precedent and to do so would be unpopular and thus might compel some Justices to write a concurring opinion to add their own personal explanation.[124] We also control for whether the petitioner won the case because the Supreme Court is prone to reverse cases at a much higher rate.

We also include a categorical variable for whether the Court declares a federal, state, or local law as unconstitutional or not. Because it is categorical, we make "no declaration" the omitted baseline condition. We also include a variable to control for the cases that have a lower court dissent. We do this because lower court dissents can signal clear differences in a case and expose fault-lines, making the likelihood of writing a concurrence greater. We also control for whether there was an unusual disposition in the case on the basis that Justices are less likely to "silently" join a majority opinion without writing a concurring opinion if that majority disposition does something out of the ordinary. Finally, we include fixed effects for each Justice and also each term. These fixed effects allow us to be confident that any idiosyncratic or stylistic effects in the opinions, or any differences due to a particular term or natural court that do not represent expertise will be captured by those coefficients. The fixed effects ensure that our expertise estimates will be indicative of expertise in general and not just a handful of Justices.

Table 2 contains the model estimates for both criminal procedure and economic activity for both regular and special concurrences. The results largely support our expectations with some important exceptions. Specifically, looking first at the left hand side of Table 2—the two models for regular concurrences—we see that

---

123.   We also estimated the models as poisson regressions for count dependent variables and with ordinary least squares regression, and both strategies return very similar results.
124.   Collins, *supra* note 117, at 156.

the "amount" variable is statistically significant and signed negatively for both criminal procedure and economic activity. This indicates that as relative expertise increases, the number of regular concurring opinions decreases. For the "structure" variable, it is significant and negatively signed for criminal procedure, but it is not significant for economic activity.

We are not sure why expertise in structuring knowledge would matter less in the economic activity domain, but perhaps it is possible that the two legal areas of the law are sufficiently different, with economic activity not requiring a complex structure to store the domain knowledge. This seems plausible given that the issue domain of criminal procedure is much younger, starting to rapidly increase in size in the 1940s and 1950s after the famous footnote four in *United States v. Carolene Products, Co.*,[125] and the issue area has been a near-constant source of contention in American politics ever since then.[126] In contrast, the body of law governing economics has been around for much longer, undergoing relatively less change compared to criminal procedure. However, this is just speculation.

---

125. 304 U.S. 144, 152 n.4 (1938).
126. *See* WARD & WEIDEN, *supra* note 69, at 203.

Table 2

DOES EXPERTISE EXPLAIN THE NUMBER OF CONCURRING OPINIONS?

	Regular Concurrences		Special Concurrences	
	Criminal Procedure	Economic Activity	Criminal Procedure	Economic Activity
Expertise: Amount	-.360** (.069)	-.352** (.120)	-.137* (.074)	-.082 (.114)
Expertise: Structure	-.197** (.070)	.012 (.094)	.020 (.083)	.037 (.077)
Precedent Altered	.523** (.184)	-.247 (.370)	.283 (.208)	.354 (.320)
Petitioner Win	-.027 (.099)	.076 (.139)	.068 (.100)	-.032 (.130)
Federal Law Declared Unconstitutional	.250 (.334)	-14.51** (.840)	-.613 (.910)	1.93** (.647)
State Law Declared Unconstitutional	-.050 (.227)	.224 (.292)	.616** (.184)	-.200 (.302)
Local Law Declared Unconstitutional	-1.24** (.448)	1.09** (.433)	-19.62** (.633)	1.44** (.338)
Lower Court Dissent	-.166 (.108)	.228 (.159)	-.013 (.074)	-.101 (.168)
Unusual Disposition	.872** (.372)	.110 (.465)	-.252 (1.01)	.032 (.366)
Justice Fixed Effects	Yes	Yes	Yes	Yes
Term Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	1452	1425	1452	1425

Note: *=p<.05, **=p<.01, one tailed test. Cell entries for each model represent negative binomial regression coefficients with robust standard errors in parentheses. The dependent variable in all models is the number of concurrences written, ranging

from 0 to 4. Both models contain fixed effects controlling for each individual Justice and each term/year, but are not shown due to space constraints. The three variables for declaring a law unconstitutional are categorical relative to an omitted baseline of no law declared unconstitutional.

Shifting to the right side of Table 2, focusing on special concurrences, we see that only the amount variable for criminal procedure cases is significant. What this means is that judicial expertise appears to have relatively little effect on reducing the number of special concurrences written. This may be intuitive in some sense given that if a Justice disagrees with the rationale of the majority opinion, that disagreement may have roots in the Justices' ideology or belief about how the law should be, and so no amount of expertise in the form of framing or addressing of counter-arguments will satisfy the concurring Justice.

To help understand the substantive meaning of the coefficients, we used the model output from the criminal procedure cases for regular concurrences to generate expected counts. Holding other variables at their mean, we varied the amount of knowledge variable from two standard deviations below the mean to two standard deviations above the mean, and we are able to use the model to predict the number of expected regular concurrences. We find that at two standard deviations below the mean, the expected count of regular concurrences is .35. At one standard deviation below the mean, the expected count is .24; at the mean the expected count is .17; at one standard deviation above the mean it is .12; and at two standard deviations above the mean it is .08. These changes may seem small at first glance, especially since we would not even expect a concurring opinion with the least expert Justice writing the majority opinion. However, the shifts in these values are actually quite large considering the fact that by far the modal outcome is 0 concurring opinions.

Thus, to better understand where expertise can have an important impact, consider a case where all other values are at their mean (like the example above), but the Supreme Court is altering a precedent and issuing an unusual disposition (two factors shown to increase the likelihood of the number of regular concurrences in Table 2). When we perform the same sort of simulation where we vary only the amount of expertise, we see the expected counts change from: 1.35 (-2 s.d.), .94 (-1 s.d.), .66 (mean), .46 (+1 s.d.), to .32 (+2 s.d.). Thus, in a case where the other conditions would highly predict a concurring opinion being written if a novice were writing the majority opinion, a highly skilled expert writing the majority

opinion would likely be able to prevent that concurrence from being written.

To ensure that our results are not just limited to one dependent variable, we also applied our expertise measures to see whether they predicted the size of the majority coalition. Without going into great detail, it should be sufficient to know that both the amount and structure of expertise significantly predicted the size of the coalition for criminal procedure cases, and the amount variable was significant for the economic activity issue area. Thus, our results are robust across other outcome measures of importance.

CONCLUSION AND DISCUSSION

In summary, we highlighted how expertise is dynamic and that it is crucial to conceptualize it this way when the legal environment is rapidly becoming more complex. We then highlighted how previous attempts to measure judicial expertise quantitatively, which occurred primarily in the political science literature, all attempted to equate experience with expertise—something that is problematic in a dynamic domain. Next, we outlined approaches to measuring expertise in psychology as well as network science. We showed that we can use text written by experts to construct networks with properties that have shown in prior research to indicate expertise. We then examined the underlying structure of these indicators to reveal two facets to expertise that were consistent with prior research.

Our findings should spark great interest in how Supreme Court opinions are written. We found that not all Justices developed expertise as they gained more experience. In fact, some Justices increased their expertise, but others either stayed the same or declined. This finding meshed with our expectation, given that the legal domain is becoming increasingly complex. We think it is fascinating and noteworthy that even in the face of this growing complexity, some Justices were able to become greater experts in areas of the law. Unfortunately, there are also other normative consequences that our findings suggest. In particular, the finding that some Justices' expertise declined over time (for whatever reason) suggests that legal policy may have suffered. While we do not know the causes of why expertise increases or decreases, and that was not the purpose of this paper, our findings suggest that expertise is not a simple concept, but rather a multifaceted one that may be changing as the environment changes. Furthermore, this can be complicated by the fact that Justices choose to adopt a legal philosophy or style of

jurisprudence to try and become more principled or consistent in their decision making. What we should keep in mind, however, and perhaps most importantly, is that our measures of expertise significantly predicted the number of concurring opinions written in an expected way. Thus, we can be confident that judicial expertise plays an important role in crafting legal policy and how a case decision influences its relevant audiences.

In this paper we have focused on how to measure judicial expertise and only one of its implications. However, it is important to understand that developing a new measure of expertise is vital because it becomes possible to answer so many other important research questions. For example, in future work we might investigate whether experts are better able to ensure compliance with their decisions by whether lower court judges are more likely to treat expertly written opinions favorably. Being able to address these sorts of questions, we would then be able to speak more forcefully about important substantive implications about how expertise influences legal policy. As we saw above, this can come in the form of an extra concurring opinion being written that may constrain or limit the influence of a majority opinion. This lack of clarity might then lead to more litigation and conflict in the lower courts, something that Justices generally want to avoid. In sum, understanding that judicial expertise is dynamic is crucial to having a judiciary that is not only independent but also able to pull its own weight in a separation of powers system.

Finally, this research extends the use of complex networks on texts into the legal academy. The legal field is not immune to changes that are already happening in other areas of society, where extracting information from text has already become increasingly important. With our approach here, we were able to use common topological properties of co-occurrence networks and extract information relevant to expertise. Importantly, however, this could be easily extended to other network measures that may be correlated with expertise. One possibility is the "small world" property of networks,[127] where the degree of "small worldness" is thought to reflect ease of mental navigation.[128] Furthermore, one could also

---

127.   Duncan J. Watts & Steven H. Strogatz, *Collective Dynamics of 'Small-World' Networks*, 393 NATURE 440 (1998).

128.   Ramon Ferrer i Cancho & Richard V. Solé, *The Small World of Human Language*, 268 PROC. ROYAL SOC'Y LONDON B 2261 (2001); Michael S. Vitevitch, *What Can Graph Theory Tell Us About Word Learning and Lexical Retrieval?*, 51 J. SPEECH, LANGUAGE & HEARING RES. 408, 414 (2008).

incorporate semantic relationships between text concepts as a way to supplement the reliance on topological properties and tap into the rich semantic meaning of the texts. Regardless, we believe our approach to measuring judicial expertise is robust, and it should open the door for future legal researchers who want to model human knowledge of the law with text.