University of Toronto Law Journal

Spring, 2013

Article

*159 Opinion Writing and Authorship on the Supreme Court of Canada [FNd1]

Kelly Bodwin, Jeffrey S Rosenthal, Albert H Yoon [FNa1]

Copyright © 2013 by University of Toronto Press; Kelly Bodwin, Jeffrey S Rosenthal, Albert H Yoon

In contrast to other branches of government, the Supreme Court of Canada operates with relatively lean staffing. For most of the Court's history, its justices alone determined which cases to review, heard oral argument, and wrote opinions. Only since 1967 have justices have been aided in these responsibilities by law clerks. While interest abounds in the relationship between justices and their clerks -- particularly the writing of opinions -- very little is known. This article analyses the text of the Court's opinions to better understand judicial authorship. We find that justices have distinct writing styles, allowing us to distinguish them from one another. Their writing styles also provide insight into how clerks influence the writing of opinions. Most justices in the modern era possess a more variable writing style than their predecessors did, both within and across years, providing strong evidence that clerks are increasingly involved in the writing of judicial opinions.

Keywords: Supreme Court of Canada, judicial authorship, textual analysis, **stylometry**, law clerks

I *Introduction*

'[A]mbitious judges realize that law clerks help them attain their ambitions. People want the best ghostwriters.'

--Judge Richard Posner [FN1]

The Supreme Court of Canada stands in stark contrast to Parliament. By comparison to members of Parliament, who have legions of staff to help them execute their duties, the Court is relatively parsimonious. With the aid of only their law clerks, the justices are charged with the ***160** responsibility of deciding which among hundreds of applications for leave to appeal to hear and of producing written opinions for every judgment. Each justice is authorized to hire no more than three clerks, [FN2] typically in their twenties [FN3] and recent graduates from elite [FN4] Canadian law schools.

While scholars have long been fascinated with the Court generally, [FN5] relatively little, if anything, has been written about the subject of authorship in the Supreme Court of Canada, which stands in sharp contrast to the United States, where there is a rich vein of anecdotal accounts as to which justices of the Supreme Count did or did not write their own opinions. [FN6] For example, in the United States, Justices Holmes, Cardozo, and Douglas, by reputation, wrote their own opinions. [FN7] Later justices, such as Justice Marshall, were reputed to have relied more on their clerks. [FN8] These anecdotal accounts of authorship are strongly supported by recent statistical analyses of the judges' own writing. [FN9]

Relatively little is known about the role of judicial clerks in the Supreme Court of Canada. Some scholars assert that

judicial clerks in ***161** Canada have grown in power, rivalling those at the United States Supreme Court. [FN10] Clerks not only assist in evaluating appeals and writing bench memos, they are alleged to have created new judicial doctrines on behalf of their justices. [FN11] Others contend that the justices do their own writing. [FN12] Lorne Sossin, Dean of Osgoode Hall School of Law and a former Supreme Court clerk, acknowledges that 'clerks now play a more significant role in the institution of the Supreme Court than ever before' but maintains that '[c]lerks do what they are told to do and have no independent duties or tasks beyond the instructions they receive from the Justices.' [FN13] Other scholars echo this viewpoint, having written that justices 'have not surrendered considerable discretion and initiative to their own clerks.' [FN14] While recent scholarship provides evidence that older workers may be more productive than their junior co-workers, [FN15] the Court is unique in that it places a heavy workload on a relatively small number of older individuals.

Given the prominence of the Court, the importance of judicial authorship extends beyond mere academic inquiry. First, the Court provides the final say on judicial matters in Canada. Second, in a common law regime, where judicial precedent matters, the reasoning that the Court applies in an opinion is arguably just as important as the question of which party prevails. Lower courts, practitioners, and legal scholars care-fully deconstruct the reasoning of the Court. For that reason, the extent ***162** to which the reasoning of the Court reflects the justices' thinking -- or those of their clerks -- matters.

At the same time, the institutional design of the Court provides a natural identification strategy to test our hypothesis. Supreme Court clerk-ships are typically for a single term, running from August or September for a full year. If justices indeed depend heavily on their clerks, the annual turnover in judicial clerks should reveal this reliance. Our premise is that justices who rely more on their clerks in the opinion-writing process possess a more variable writing style than their less reliant colleagues, both within and across years. Moreover, the relatively recent inception of judicial clerks on the Court in 1967 suggests that the writing variability in the pre-clerk years should be lower than the post-clerk years. The progression toward multiple clerks within each chambers should, for those justices relying heavily on their clerks in the opinion-writing process, further increase their writing variability.

In this article, we analyse the text of majority opinions written by all justices on the Court. We construct a variability score based on the justices' own use of common function words (e.g. *for, have, with*). We find that each justice possesses a unique variability score. More importantly, however, even justices with nearly identical variability scores possess distinct writing styles based on these function words. These differences, in turn, allow us to correctly identify authorship in a pairwise comparison of justices. We find that current and recent justices possess higher variability scores than the justices preceding them, providing statistical evidence that justices, on average, rely more on their clerks when writing opinions.

This article proceeds as follows. We describe in Part II our methodology of measuring writing variability through the use of common function words (e.g., *an, have, what*). In Part III we describe our data on the Court's decisions. We report our main results in Part IV, showing the justices' variability scores both within and across cohorts and our ability to use these scores to accurately predict authorship in pairwise comparisons of justices. We discuss the implications of our results in Part V.

II *Methodology*

In analysing the text of judicial opinions from the Supreme Court of Canada, we follow the methodology established by Rosenthal and Yoon in their analysis of the United States Supreme Court. [FN16] We briefly summarize the approach here.

*163 Judicial writing, and writing in general, is influenced on two factors. The first is the subject or topic of the writing; the second is the author's writing style -- that is, her word choice (diction) and sentence structure (syntax). Our statistical analysis focuses on the latter factor, falling within the broader discipline of **stylometry**. [FN17] Specifically, we analyse commonly used words, often referred to as function words. [FN18] Our central assumption is that an author's use of such words operates independently of the subject or. topic of the writing. [FN19]

 The intuition behind our approach -- supported by recent scholarship -- is that individuals' writing styles tend to be fairly consistent; thus, greater variability in writing style reflects that a justice delegates more of her writing responsibilities to her law clerks: the more people actively involved in writing judicial opinions, all things equal, the more variable the writing style. For example, if Justice A uses the word 'this' about 2% of the time and the word 'some' about 3% of the time, then opinions authored solely by Justice A will tend to follow these percentages fairly consistently and thus have fairly low variability scores. Similarly, if Justice B uses the word 'this' about 1% of the time and the word 'some' about 4% of the time, then opinions authored solely by Justice B will follow these percentages fairly consistently and again have fairly low variability scores. In particular, solely authored items are characterized by low *variability* of word frequencies, not by the specific word frequencies them-selves.

On the other hand, if Justice A instead lets many different law clerks write her different opinions, then some of those law clerks might use 'this' much more than 2% of the time, others might use 'this' much less than 2% of the time, and still others might use 'some' much more than 3% of the time, and so on. If so, then the word frequencies of Justice A's opinions will change significantly from one opinion to the next and they will thus have much higher variability scores. (Of course, it is conceivable that some of the law clerks will happen to have very similar word usage to that of Justice A herself, but it is highly unlikely that many of the clerks ***164** will be very similar in all 63 function words that we consider.) In this way, high variability scores can potentially detect cases where different law clerks are writing different opinions.

We used Mosteller & Wallace's [FN20] original list of 70 function words and deleted seven words rarely used by the justices. [FN21] The remaining 63 words appear in Table 1.

Table 1: **Stylometry** of Supreme Court text

	Sixty-three Function Words
(1-12)	*a, all, also, an, and, any, are, as, at, be, been, but*
(13-24)	*by, can, do, down, even, for, from, had, has, have, her, his*
(25-36)	*if, in, into, is, it, its, may, more, must, no, not, now*
(37-48)	*of, on, one, only, or, our, so, some, such, than, that, the*
(49-60)	*their, then, there, things, this, to, up, was, were, what, when, which*

(61-3)   *who, with, would*

For each opinion, we kept count of each of these words. Given this approach, we adopted a chi-square approach to test the distribution of the word counts we observed against a null hypothesis that, for each justice, the total count of each function word is equally likely to occur in any of the total number of her opinions.

The chi-square statistic is the following:

$$\textit{chisq} = \sum_{i=1}^{K} \sum_{j=0}^{63} \frac{(c_{ij} - e_{ij})}{e_{ij}}$$

where j is the function words numbered from $j = 1$ to $j = 63$; K represents the total number of opinions a justice has written in our data set, numbered from $i = 1$ to $i = K$; c_{ij} represents the number of times that function word j appears in opinion i; the term e_{ij} is the expected number of times that function word j would appear in judgment i; and w_i is the total number of words in judgment i.   We calculate our variability score by the following:

$$\textit{Variability Score} = \frac{\textit{chisq}}{df} = \frac{\textit{chisq}}{62 \, (K-1)}$$

where df is the degrees of freedom.   This variability score is the foundation of all our subsequent analysis.

*165 To test the null hypothesis that each justice's writing style has a uniform and random distribution of function words, we randomly generated 200 pseudo-documents, each consisting of 2 000 independently and randomly generated words. We set up the documents such that each of the 2 000 words had a 70% probability of being a non-function word and a 30% probability of being a function word from Table 1. The score approximated 1 under the null hypothesis. We repeated this experiment 10 times, which produced a mean variability score of 1.004622, with a standard deviation of 0.001702, consistent with a null hypothesis with a true mean equal to one. A variability score that exceeds 1 reflects greater variability. In other words, the higher the variability score, the more variable the judge's writing style, based on the function words in Table 1.

Finally, because chi-square values are less stable when expected cell counts for given words approximate zero, we excluded function words with a very low frequency, as well as opinions shorter than 250 words.

Our central identification strategy for discerning writing variability was law clerks themselves. Law clerks typically serve for only one year and are replaced annually by a new set of law clerks. If law Clerks vary in their own writing style, then justices who rely more on their clerks for drafting or writing of opinions would reveal greater variability in writing style than justices who do their own writing.

While we believe that this identification strategy provides a compelling proxy for how much justices rely on their law clerks, we recognize that it is ultimately untestable. An alternative explanation would always exist: even in a world where all justices wrote their own opinions, some justices would have more variable writing styles than others. The relationship between justices and their clerks remains a closely guarded secret in the United States [FN22] and even more so in Canada. But we maintain that greater variability in writing both within and across years strongly suggests greater reliance by justices on their law clerks. As an additional validity check, two prominent United States Courts of Appeal judges widely reputed to write their own opinions, Richard A Posner and Frank Easter-brook, exhibit variability measures that were both markedly lower and less variable than those of contemporary United States justices. [FN23]

*166 III *Data*

We constructed a data set of the written judgments from the Supreme Court of Canada, available on the Lexum Web site. [FN24] This Web site -- a collaborative effort of the Court, Lexum, and the University of Montreal -- includes all published opinions since the inception of the Court, dating back to 1876. While the data set spans the entire history of the Court -- comprising 74 justices [FN25] -- our analysis focuses on justices in the period after 1900.

We analysed the opinions using software in C and Unix that down-loaded the decisions directly from the Lexum Web site. The opinions for each case were contained in HyperText Markup Language (HTML) format. [FN26] Specifically, the program converts the HTML pages into plain text and extracts from each opinion text not written by the justices them-selves, including headnotes and procedural synopses (though quotations within the text are not removed). Our program also separates majority from concurring and dissenting opinions. The resulting text files were then spot checked to ensure accuracy and avoid downloading errors.

Although we analysed each type of opinion within each judgment -- majority, concurring, dissenting -- we focused on majority opinions. Concurring and dissenting opinions have become more common, but most judgments today consist of a single, majority opinion. Because dissents and concurrences are often shorter than majority opinions, methodologically they create instability in the textual analysis. We also excluded *per curiam* opinions and co-authored opinions, both sparingly used by the Court. [FN27]

IV *Results*

Our analysis of the results builds from our variability score. We first established that the justices' variability scores are statistically distinguishable from the null hypothesis. We then constructed a bootstrap approach to determine whether, in a pairwise comparison, justices' ***167** variability scores were statistically distinguishable from one another. To test whether our variability score could accurately predict authorship, we designed a linear classifier, enabling us to compare scores for each justice, looking in particular at the beginning and end of their tenure. Third, we tested the accuracy of the linear classifier through a 'leave-one-out cross-validation.' As mentioned earlier, our analysis, unless otherwise stated, is based on majority opinions.

A VARIABILITY SCORES

We begin by describing the variability scores of the justices on the Supreme Court of Canada. These scores provide a measure of the each justice's writing style with which we can compare one to another. Table 2 lists the 38 justices who joined the court after 1948, including the province they represent, the prime minister who appointed them, when they joined and left the Court, how many majority opinions they wrote, the average length of their majority opinions, and im-

portantly, their variability score.

A higher than average variability score, as noted earlier, is consistent with the view that the given justice has relied more on her law clerks in writing opinions. It bears repeating that we acknowledge the possibility of an alternative explanation that, even in a world where justices did their own writing, some justices would have a greater variability in writing style than others. It is also possible that, even in a world where all justices relied on clerks to write opinions, some clerks would be better than others at mimicking their justices' writing styles. Thus, the variability measure could be capturing differences in clerks' abilities. Because of the incompatibility of the textual analysis, in this table and throughout our analysis, we exclude justices who wrote their opinions in French. [FN28]

The variability scores range from a high of 5.02 for Justice Frank Iacobucci to a low of 2.62 for Justice Louis-Philippe de Grandpré. The variability score for each justice is statistically significantly larger than the null hypothesis of uniformly and randomly distributed function words. As an example, Justice Louis LeBel's variability score of 4.43 reflects a chi-square statistic of $4.43 \times 63 \times (114-1) = 31,537.17$. The null hypothesis has a chi-square statistic of $63 * (114-1) = 7,119$. The p value corresponding to Justice LeBel's score is less than 0.00001, allowing us to reject the null hypothesis. For close followers -- and perhaps current and former law clerks -- of the Court, the variability scores may support or contradict their beliefs about the extent to which justices write their own opinions.

Num	Justice	Province	Appointing Prime Minister	Year Joined Court	Year Left Court	Years On Court	Opinions	Average Word Length	Variability Score
1	Thomas Cromwell	NS	Harper	2008	2011	3	14	5 748	4.75
2	Marshall Rothstein	MB	Harper	2006	2011	5	39	6 551	4.60
3	Louise Charron	ON	Martin	2004	2011	7	45	6 312	4.10
4	Rosalie Abella	ON	Martin	2004	2011	7	52	4 844	4.27
5	Morris J. Fish	QC	Chrétien	2003	2011	8	55	3 758	3.

									00
6	Marie Deschamps	QC	Chrétien	2002	2011	9	48	7 371	4.25
7	Louis LeBel	QC	Chrétien	2000	2011	11	115	7 367	4.43
8	Louise Arbour	ON	Chrétien	1999	2004	5	47	7194	4.13
9	William Ian Corneil Binnie	ON	Chrétien	1998	2011	13	109	7 733	3.79
10	Michel Bastarache	NB	Chrétien	1997	2008	11	81	7 748	4.41
11	John C. Major	AB	Mulroney	1992	2005	13	99	5 161	4.33
12	Frank Iacobucci	ON	Mulroney	1991	2004	13	123	8 783	5.02
13	William Stevenson	AB	Mulroney	1990	1992	2	14	3 108	2.84
14	Beverley McLach-lin	BC	Mulroney	1989	2011	22	290	5 439	4.56
15	Peter Cory	ON	Mulroney	1989	1999	10	121	6 173	4.20

16	Charles Gonthier	QC	Mulroney	1989	2003	14	102	7 459	4. 94
17	John Sopinka	SK	Mulroney	1988	1997	9	145	4 476	3. 17
18	Claire L'Heureux-Dubé	QC	Mulroney	1987	2002	15	68	6 496	4. 24
19	Gérard La Forest	NB	Mulroney	1985	1997	12	141	6 721	3. 77
20	Gerald Eric Le Dain	ON	Trudeau	1984	1988	4	31	5 531	3. 15
21	Bertha Wilson	ON	Trudeau	1982	1991	9	112	5 736	3. 97
22	Antonio Lamer	QC	Trudeau	1980	2000	20	228	5 542	4. 32
23	Julien Chouinard	QC	Clark	1979	1987	8	56	4 234	3. 14
24	William Rogers McIntyre	BC	Trudeau	1979	1989	10	108	4 068	3. 16
25	Yves Pratte	QC	Trudeau	1977	1979	2	24	4 436	3. 10
26	Louis-Philippe de	QC	Trudeau	1974	1977	3	71	2 435	2.

	Grandpré							62
27	Jean Beetz	QC	Trudeau	1974	1988	14	82	6 807
								4. 47
28	Robert George Brian Dickson	MB	Trudeau	1973	1990	17	213	5 432
								3. 72
29	Bora Laskin	ON	Trudeau	1970	1984	14	245	3 282
								3. 04
30	Louis-Philippe Pigeon	QC	Pearson	1967	1980	13	206	2 727
								2. 78
31	Wishart Flett Spence	ON	Pearson	1963	1978	15	166	3 272
								3. 19
32	Emmett Matthew Hall	SK	Diefenbaker	1962	1973	11	94	2 456
								2. 98
33	Roland Almon Ritchie	NS	Diefenbaker	1959	1984	25	286	2 856
								2. 81
34	Wilfred Judson	ON	Diefenbaker	1958	1977	19	211	1 708
								2. 62
35	Ronald Martland	AB	Diefenbaker	1958	1982	24	274	2 999
								3. 81
36	Henry Grattan Nolan	AB	St. Laurent	1956	1957	1	9	2 950
								2. 69

| 37 | Douglas Charles Abbott | QC | St. Laurent | 1954 | 1973 | 19 | 121 | 1 183 | 2. 73 |
| 38 | John Robert Cartwright | ON | St. Laurent | 1949 | 1970 | 21 | 271 | 2 280 | 2. 82 |

Note: Table excludes justices who wrote their opinions in French. Justices are arranged in reverse chronological order, by year of appointment. Justices who were appointed to the court prior to 1949 are listed in the Appendix, Table A1.

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

*170 Collectively, Table 2 suggests that the variability in writing has increased over time across justices.

Figure 1, showing variability scores for justices beginning in 1900, graphically shows this upward trend. The sloped dotted line represents the regression line over the entire period, with a significant positive slope ($p > 0.001$). Although justices sitting together differed in their writing variability, the regression line is positively sloped over time.

Figure 1: Variability scores over time

Note: Sloped dotted line represents regression line over entire period. Sloped solid lines represent regression line within each time period.

It is worth noting the differences in slope before and after the inception of law clerks in 1967, as represented by the thick black line in Figure 1. For the period 1900-66, the regression line is flatter relative to the regression line for the entire period, with variability scores increasing by about 0.15 per year. Conversely, for the period 1967-2010, the ***171** regression line is steeper relative to the overall regression line, increasing at a rate of 0.33 per year. [FN29] Interestingly, attempts to fit a jump discontinuity at 1967 (as opposed to the kinked regression in Figure 1) did not lead to a statistically significantly better fit. This means there was no immediate change in variability scores with the advent of clerks; rather, one might infer that justices gradually adapted to the presence of clerks and their variability scores increased more quickly as a result. It is also of note that no significant jump or change in slope was found at 1989, [FN30] the year of expansion to a three-clerk system. Perhaps, then, the inflation increase in the number of clerks was not an exogenous shock, but rather a response to the justices' evolving reliance upon their assistants.

Similarly, Table 3 shows an upward trend in writing variability when categorizing time periods by the presiding chief justice. This evidence illustrates that justices' writing style was less variable prior to law clerks and became increasingly more variable following the introduction of law clerks.

Table 3: Variability score by chief justice

Justice	Province	Elevating Prime Minister	First Year as Chief	Last Year as Chief	Total Years on Court	Total Opinions	Average Word Length	Variability Score

Beverley McLach-lin	BC	Chretien	2000	present	22	807	6 788	4.61
Antonio Lamer	QC	Mulroney	1990	2000	20	- 1 194	6 190	4.36
Robert George Bri-an Dickson	MB	Trudeau	1984	1990	17	541	5 398	3.79
Bora Laskin	ON	Trudeau	1973	1984	14	1 179	3 563	3.43
Gérard Fauteux	QC	Trudeau	1970	1973	24	392	2 502	3.06
John Robert Cartwright	ON	Pearson	1967	1970	21	354	2 290	3.18
Robert Taschereau	QC	Pearson	1963	1967	27	383	2 195	3.19
Patrick Kerwin	ON	St. Laurent	1954	1963	28	854	2 100	3.04
Thibaudeau Rinfret	QC	King	1944	1954	30	605	2 182	2.96
Sir Ly-man Poore Duff	BC	Bennett	1933	1944	38	232	2 256	3.01
Francis Alexander Anglin	ON	King	1924	1933	24	192	2 063	2.76
Sir Louis Henry Dav-	PEI	Borden	1918	1924	23	216	1 497	2.39

ies

Sir Charles Fitzpatrick	QC	Laurier	1906	1918	12	305	1682	2.28
Sir Henri Elzéar Taschereau	QC	Laurier	1902	1906	28	105	1536	2.43
Sir Samuel Henry Strong	ON	Thompson	1892	1902	27	204	2 028	3.01
Sir William Johnstone Ritchie	NB	Macdonald	. 1879	1892	17	255	2 557	3.35
Sir William Buell Richards	ON	NA	1875	1879	4	49	3 200	4.08

Notes: Variability (V) analysis not performed for francophone justices. Sir Williams B Richards was appointed as chief justice upon creation of the Court. Table excludes justices who wrote their opinions in French.

These aggregate scores, while informative, are incomplete. An aggregate score, by itself, does not reveal how the justices' scores may vary over their tenure on the Court. Theoretically, there are reasons to suggest both how they may increase and decrease. The job is intellectually demanding and may impose a steep learning curve at the beginning. Also, as justices become older, their willingness to delegate writing obligations may increase.

Figure 2 gives a sample of recent justices throughout their tenure. Justice La Forest, for example, with an aggregate variability score of 3.77, had lower -- and remarkably consistent -- writing variability during his first five years on the Court. His writing style subsequently became more variable, both within and across years. Justices Lamer, L'Heureux-Dubé, Sopinka, Cory, McLachlin, and Iacobucci similarly showed an increasing variability in writing style during their tenures. Justice Gonthier, by contrast, was one of the few justices whose writing style became less variable over the bulk of his tenure, while Justice Major's writing style steadily grew more variable during his first eight years on the Court but became less variable during his final five years on the Court.

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

Figure 2: Year-to-year Variability scores, select justices

As stated in the methodology section, law clerks serve as the 'central identification strategy' for discerning writing variability. One limitation of this approach is that post-1967, all justices are reputed to have hired law clerks, meaning that there is no reason to expect variation across justices subsequent to 1967. Justice Judson provides an exception to the ***173** institutional trend. For the 1974--5 term, Justice Judson did not hire a law clerk, [FN31] meaning that he wrote his own opinions. Figure 3 supports the view that the presence and absence of a law clerk affects writing variability. Following the implementation of law clerks, Justice Judson's writing variability increased in the short term (1969-70 through 1971-2), before declining. While one should exercise caution against over-interpretation of these results, given the variation over time, Justice Judson's variability score in the 1974--5 term was the lowest of his career. Moreover, for the 1975-6 term, after Justice Judson hired a law clerk, his writing variability sharply increased.

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

Figure 3: Variability score, Justice Judson

Note: The square marker indicates year in which Justice Judson did not have a law clerk.

B SIGNIFICANCE TESTS

Our variability measure for each of the justices allows us to reject the null hypothesis that each justice's writing style follow a uniform and random distribution of function words. It does not, however, speak to whether the differences in variability scores from one justice to another are statistically meaningful. For this question, no simple analytic test exists. Because we rejected the null in each instance, the justices' scores by ***174** definition do not follow a chi-square distribution. To complicate matters, it is not possible to determine analytically the shape of the distribution.

To get around this analytical intractability, we used a bootstrap test: repeated sampling with replacement from a given sample, [FN32] which enables us to determine the distribution empirically. We randomly selected for each justice one-hundred authored majority opinions, with repetition. [FN33] We computed the variability score for this sample. We then repeated the process 1 000 times for each justice, generating 1 000 different possible variability scores, depending on the opinions selected for each sample.

We then did pairwise comparisons of justices, based on their bootstrap variability scores. This process created 1 million (1 000 x 1 000) pairs of variability scores. We then tallied the number of pairs where the ***175** variability score was greater for Justice A than for Justice B. The resulting fraction provides an estimate of the probability that Justice A's variability score is greater than Justice B's variability score for a random selection of judgments. From these pairings, we can estimate the distribution function for the difference in variability scores between the two justices, from which we can compute a 95% confidence interval (CI). As with any CI, a wholly positive or negative CI indicates that the difference in variability scores between Justices A and B is statistically meaningful.

Table 4 illustrates two examples of bootstrap pairwise comparisons. Each comparison reports the bootstrap scores for each of the two justices. The bootstrap variability score is similar to the observed variability scores fore each justice in Table 2. The greatest divergence is for Justice L'Heureux-Dubé, in large part because she wrote only 68 majority opin-

ions, far fewer than the other three justices in Table 4. Unsurprisingly, sampling one-hundred opinions from a pool of sixty-eight opinions will generate greater divergence from the observed variability score. For the first comparison, Justice McLachlin had a higher bootstrap variability than Justice L'Heureux-Dubé. A probability less than 0.05 or greater than 0.95 indicates a statistically significant difference in the bootstrap variability score. The probability that Justice McLachlin's variability was less relative to that of Justice L'Heureux-Dubé was 0.09, which is not statistically significant (at the $p < 0.05$ level). Conversely, Justice Iacobucci's higher bootstrap score relative to Justice Sopinka's was statistically significant.

Table 4: Variability score pairwise bootstrap - Sample comparisons

Bootstrap V4 score (McLachlin)	Bootstrap V score (L'Heureux-Dubé)	P(McLachlin < L'Heureux-Dubé)	95% CI: (McLachlin-L'Heureux-Dubé)
4.50	3.91	0.0868	(-1.6236,0.2139)

Bootstrap V4 score (Iacobucci)	Bootstrap V score (Sopinka)	P(Iacobucci < Sopinka)	95% CI: (Iacobucci-Sopinka)
4.94	3.12	0.0000	(-2.4715,-1.1360)

Table 5 reports the pairwise comparison of all justices from the 1992-7 'Lamer natural Court' period (i.e., a period when Justice Lamer was chief justice and no other justices joined or left the court). The original variability score for each justice is on the left-most column. [FN34] Each cell reports the probability that the justice listed in the row header has a **177** lower variability score than the justice listed at the column header. For example, the probability that Justice McLachlin has a lower variability score than Justice La Forest is 0.0282, a statistically significant difference. Across the shaded diagonal, the variability scores are mirror images, meaning that the probability that Justice La Forest has a lower variability score that Justice La Forest is 0.9718.

[Note: The following table/form is too wide to be printed on a single page. For meaningful review of its contents the table must be assembled with part numbers in ascending order from left to right. Row numbers, which are not part of the original data, have been added in the margins and can be used to align rows across the parts.]

****
************** This is piece: 1
****

	V4 Score	Antonio Lamer	Gerard La Forest	Claire L'Heureux-Dubé	John Sopinka	Charles Gonthier	Peter Cory	Beverley McLachlin	Frank Iacobucci
Antonio Lamer	4.32	_____	0.0764	0.3032	0.0007	0.9269	0.3727	0.6835	0.9457
Gérard La Forest	3.77	0.9236	_____	0.8255	0.0270	0.9998	0.8871	0.9718	0.9998
Claire L'Heureux-Dubé	4.07	0.6968	0.1745	_____	0.0009	0.9916	0.6124	0.8648	0.9956
John Sopinka	3.17	0.9993	0.9730	0.9991	_____	1.0000	0.9997	1.0000	1.0000
Charles Gonthier	4.94	0.0731	0.0002	0.0084	0.0000	_____	0.0188	0.2316	0.6013
Peter Cory	4.20	0.6273	0.1129	0.3876	0.0003	0.9812	_____	0.7738	0.9865
Beverley McLachlin	4.56	0.3165	0.0282	0.1352	0.0000	0.7684	0.2262	_____	0.8026
Frank Iacobucci	5.02	0.0543	0.0002	0.0044	0.0000	0.3987	0.0135	0.1974	_____
John C. Major	4.33	0.5196	0.2478	0.2478	0.0000	0.9695	0.3711	0.7148	0.9844

Table 5: Variability score pairwise bootstrap - Sitting justices during Lamer natural Court (1992-7)

Note: The variability score column represents the variability score for each justice. The remaining cells represent the probability that (variability score for row justice) < (variability score for column justice).

```
*********************************************************************
************** This is piece: 2
*********************************************************************
```

1	
2	John C. Major
3	0.4804
4	0.9566
5	0.7522
6	1.0000
7	0.0305
8	0.6289
9	0.2852
10	0.0156
11	_____
12	

As a general matter, the greater the difference in variability scores, the more likely the bootstrap will produce a statistically distinguishable difference. For example, because Justice Iacobucci had the highest variability score (5.02) of this natural Court, the bootstrap comparison of his score was statistically distinguishable from those all of the justices, save Justice McLachlin and Justice Gunthier, the two justices with the next highest variability scores. Conversely, Justice Sopinka, by a considerable margin, had the lowest variability score (3.17) in this cohort, statistically distinguishing him

from all other justices. In instances where a pair of justices have nearly identical variability scores -- as is the case with Justice Lamer (4.32) and Justice Major (4.33) -- the bootstrap probabilities converge to 0.50. Accordingly, justices with variability scores near the median of the Court were, perhaps unsurprisingly, less distinguishable than justices with relatively low or high variability scores.

Of the 36 unique justice pairings, 18 (or 50%) were statistically significant. Even when falling short of statistical significance at the $p < 0.05$ level, 30 (or 83%) of these pairings report a pairwise probability less than 0.30 or greater than 0.70.

As a separate inquiry, we constructed variability scores *within* justices in an effort to examine how individual justices may have changed over their tenure on the Court, [FN35] reported in Table 6. Political scientists studying the US Supreme Court have commented on how justices shift ideologically while on the Court. The variability measure, because it is based on common function words, should be uncorrelated with ideology. There may, nevertheless, be a connection between ideology and writing variability. To the extent that justices rely more on law clerks in writing opinions, the law clerks' may influence both the substance (ideology) and writing style.

[Note: The following table/form is too wide to be printed on a single page. For meaningful review of its contents the table must be assembled with part numbers in ascending order from left to right. Row numbers, which are not part of the original data, have been added in the margins and can be used to align rows across the parts.]

****

************** This is piece: 1

****

1	Table 6: Within-justice comparison, first five and last five years (current and past justices, post-1948)									
2			Total			First Five Years			Last Five Years		
3	Justice	Opinions	Avg Word Length	V4 Score	Opinions	Avg Word Length	V4 Score	Opinions	Avg Word Length	V4 Score
4	John Robert Cartwright	271	2 280	2.82	47	2 930	3.12	61	2 198	2.45
5	Douglas Charles Abbott	121	1 183	2.73	27	1 095	2.00	27	1 180	2.30
6	Ronald Mart-	274	2 999	3.81	53	3 145	3.75	46	3 133	3.25

	land								
7	Wilfred Judson	211	1 708	2.62	61	1973	2.62	25	1564	2.48
8	Roland Almon Ritchie	286	2 856	2.81	51	2 544	2.48	28	2 698	2.31
9	Emmett Matthew Hall	94	2 456	2.98	29	2 772	3.01	43	2 343	2.92
10	Wishart Flett Spence	166	3 272	3.19	44	2 719	2.93	52	3 488	2.98
11	Louis-Philippe Pigeon	206	2 727	2.78	50	2 165	2.54	77	3 131	2.87
12	Bora Laskin	245	3 282	3.04	81	2 623	2.98	69	3 243	3.14
13	Robert George Brian Dickson	213	5 432	3.72	63	3 130	2.89	57	7 623	4.24
14	Jean Beetz	82	6 807	4.47	30	4 364	3.70	22	9 608	4.69
15	William Rogers McIntyre	108	4 068	3.16	50	3 525	2.92	52	4 341	3.11
16	Julien Chouinard	56	4 234	3.14	36	3 885	2.98	28	4 900	3.38
17	Antonio Lamer	228	5541	4.32	36	4 529	3.83	38	8 306	4.99

18	Bertha Wilson	112	5 736	3.97	39	5 390	4.09	73	5 921	3.84
19	Gerard La Forest	141	6 722	3.77	49	4 813	2.79	49	8 027	4.11
20	Claire L'Heureux-Dubé	201	6 877	4.07	60	7 447	4.01	41	6 092	3.89
21	John Sopinka	145	4 476	3.17	73	3 616	2.60	72	5 348	3.63
22	Charles Gonthier	102	7 459	4.94	44	8 000	4.92	28	7 673	4.83
23	Peter Cory	121	6 172	4.20	60	5 110 "	3.56	52	7 089	4.40
24	Beverley McLach-lin	304	5 529	4.58	91	4 144	3.37	53	7 109	4.88
25	Frank Iacobucci	123	8 783	5.02	48	8 328	4.49	37	8 640	5.19
26	John C. Major	101	5 113	4.31	29	3 876	3.62	41	5 445	4.40
27	Michel Bastarache	84	7 794	4.46	38	7 749	4.10	36	7 812	4.89
28	William Ian Corneil Binnie	118	7 807	3.87	48	8 043	3.98	37	6 940	3.45
29	Louis	47	7 194	4.13	42	7 165	4.17	47	7 194	4.13

	e Arbour									
30	Louis LeBel	124	7 234	4.58	57	7 722	4.57	45	6 970	4.50
31	Marie Deschamps	61	7 283	4.30	25	5 931	3.70	36	8 222	4.46
32	Morris J. Fish	61	3 855	2.98	34	4 036	3.03	33	3 455	2.84
33										
34	Note: Table excludes justices serving fewer than eight years.									

```
*****************************************************************
************** This is piece: 2
*****************************************************************
```

1				
2	Bootstrapping			
3	P (First5 < Last5)	Lower 95% Confidence Interval	Upper 95% Confidence Interval	
4	0.3282	-0.96	-0.35	
5	0.8768	-0.16	0.76	
6	0.1153	-1.32	0.24	
7	0.1202	-0.53	0.13	
8	0.0208	-0.38	-0.01	

9	0.4133	-0.43	0.35
10	0.6524	-0.27	0.45
11	0.9716	-0.01	0.69
12	0.6768	-0.39	0.69
13	1.0000	0.82	1.80
14	0.9991	0.34	1.45
15	0.8237	-0.19	0.57
16	0.9755	0.00	0.72
17	0.9985	0.36	1.93
18	0.2818	-0.80	0.48
19	1.0000	0.79	1.81
20	0.3050	-0.74	0.44
21	1.0000	0.50	1.50
22	0.3246	-0.82	0.51

23	0.9957	0.19	1.40
24	1.0000	0.86	2.05
25	0.9631	-0.07	1.27
26	0.9982	0.26	1.30
27	0.9983	0.28	1.26
28	0.0143	-1.01	-0.05
29	0.4998	-0.67	0.69
30	0.4111	-0.84	0.71
31	0.9997	0.32	1.25
32	0.1021	-0.47	0.10
33			
34			

Even for those with prior judicial experience, being a Supreme Court justice is a unique experience that requires a period of adjustment. The Court's caseload, oral argument, and writing requirements differ from those of lower provincial and federal courts. An emphasis on novel issues of law likely generates new demands for most, if not all, justices. ***180** Accordingly, in Table 6 we compare justices' first five years and final five years on the Courts. We replicate the overall variability scores in Table 2 and generate variability scores for the first five and final five years, respectively. The final set of columns reports the bootstrap comparison for these two periods. We exclude justices whose total tenure on the Court was fewer than eight years.

A supermajority -- 22 out of 29, or 76% -- of justices had a higher writing variability during their final five years on the Court. This finding is consistent with the hypothesis that justices, as they reach the latter years of their tenure and the age of retirement, rely more on clerks in the drafting of their opinions. Scholars studying the US Supreme Court have speculated that older justices are prone to 'mental decrepitude,' [FN36] given that some justices have served well past the age of eighty. Because justices on the Supreme Court of Canada, like other federally appointed judges, must retire by age seventy-five, [FN37] the increased variance in their latter years is likely due, at least in part, to other factors.

One explanation for a subset of justices is an increase in administrative responsibilities. For example, four of the five chief justices listed in Table 6 had a higher writing variability in their final five years than in their first five years. For three of these four -- Chief Justices McLachlin, Lamer, and Dickson -- the increase was statistically significant (for Chief Justice Laskin, the increase was more modest and not statistically significant). The exception to this trend was Chief Justice Cartwright, whose variability actually decreased during his final five years, although this difference was not statistically significant. This general trend suggests that the administrative responsibilities that chief justices face may make it more difficult for them to write or draft their own opinions.

C IDENTIFICATION OF AUTHOR

The previous results show that the justices' writing styles are statistically distinguishable from the null and from each other. These results, in a sense, are an answer to a more modest question than the one asking whether it is possible to use function words to predict which justice authored an opinion. In our informal conversations with constitutional law scholars, the common response was that they would know the identity of the justice if the passage were well known but as a general matter they could not discern authorship from writing style, particularly of **181** function words. [FN38] To investigate our ability to determine authorship of opinions, we applied a pairwise approach (consistent with our approach for the bootstrap analysis). We compared a universe of opinions written either by Justice A or Justice B.

Intuitively, we determined authorship as follows. First, we computed typical function word frequencies for each justice, based on their other writings. Then, given an opinion of unknown authorship, we saw whether its word frequencies were closer to the typical frequencies of Justice A or to those of Justice B. Consider our previous example, in which Justice A uses 'this' 2% of the time and 'some' 3% of the time, and Justice B uses the word 'this' 1% of the time and 'some' 4% of the time. In that case, if an opinion of unknown authorship used 'this' 2.3% of the time, and 'some' 2.9% of the time, then it was probably written by Justice A. By contrast, if an opinion of unknown authorship used 'this' 0.8% of the time, and 'some' 4.1% of the time, then it was probably written by Justice B. Note that, for this author identification, we no longer concerned ourselves with frequency variability but rather with actual word frequencies, which are the most useful way of distinguishing two different writing styles. Of course, this analysis implicitly assumes that each justice has a 'typical writing style,' in spite of whatever variability -- or assistance from clerks -- may be involved. Fortunately, our results indicate that this assumption is reasonable enough and that we can identify authorship quite successfully using this approach.

To avoid over-fitting of the data, [FN39] we divided the data into two parts: training data and testing data. The training data allowed us to develop a model for classifying each judgment as being written by either Justice A or Justice B. We then used the testing data to see how well our model predicts authorship.

The accuracy of the test is based on 'leave-one-out cross-validation.' For each judgment (written by either Justice A or Justice B), this judgment serves as the test data and all other judgments (again written by either Justice A or Justice B) serve as the training data. We identify whether our model accurately attributes the test judgment to either justice. We repeat this process for all judgments written by either Justice A or Justice B, and count the number of accurate classifica-

tions.

    *182 With a linear classifier, *T* represents a training set consisting of all judgments by Justice A or B, with $|T| = n$, where *n* equals the total number of opinions. We used the linear regression model, $Y = x +$, where is an $n \times 1$ vector of independent errors with mean zero; Y re-presents an $n \times 1$ vector of: a value of -1 if Justice A actually wrote the opinion and +1 if Justice B wrote the opinion. The term *x* is an $n \times 64$ matrix, defined as:

$$x = \begin{pmatrix} 1 & f_{1,1} & f_{1,2} & \cdots & f_{1,63} \\ 1 & f_{2,1} & f_{2,2} & \cdots & f_{2,63} \\ \cdots & \cdots & \cdots & & \cdots \\ 1 & f_{n,1} & f_{n,2} & \cdots & f_{n,63} \end{pmatrix},$$

where $f_{i,j}$ are the fraction of words in judgment *i* in the training set which are from function word *j*. The least-squares estimate for <<beta>> corresponds to the maximum likelihood estimate (MLE), assuming the errors (i) to be independent and identically distributed (i.i.d.), defined as <<beta>> $= (x^T x)^{-1} x^T Y$ where <<beta>> $= ($<<beta>>$_0$, <<beta>>$_1$, <<beta>>$_n$). Given a test judgment with function words fractions of $g_1, g_2, \ldots g_{63}$, linear fit value corresponds to:

$$\text{<<Unknown Symbol>>} = \text{<<beta>>}_0 + \sum_{j=1}^{63} \text{<<beta>>}_j\, g_j$$

If <<Unknown Symbol>> < 0, we classify Justice A as the author of the test judgment; if <<Unknown Symbol>> $>= 0$ we attribute authorship to Justice B. We also tried a naïve Bayes classifier, which produces similar, but on average slightly less accurate, predictions to the linear classifier. We illustrate how the linear classifier works in Table 7, comparing two pairs of recent justices. An important note regarding interpretation: the pairwise comparison between two justices is not necessarily symmetric. In the comparison between Justice McLachlin and Justice L'Heureux-Dubé, the linear classifier accurately predicted the author of Justice McLachlin 86% of the time, and Justice L'Heureux-Dubé 80% of the time. At the same time, a pairwise comparison can be closely symmetric. The second pairwise comparison between Justice Iacobucci and Justice ***183** Sopinka produced a more symmetric result. It accurately predicted Justice Iacobucci as the author 76% of the time and Justice Sopinka as the author 75% of the time. The extent of this symmetry depends on the shape of the probability distributions of the two justices in the pairwise comparison and the degree to which they overlap. If the two distributions largely (but not completely) overlap, it leads the prediction of authorship to be significantly higher for one of the justices. [FN40]

    Table 7: Linear classifier, sample comparisons

Justice *A*	Justice *B*	Fraction Ac-	Success Rate	Fraction Ac-	Success Rate Pre-

		curately Predicting Justice *A*	Predicting Justice *A*	curately Predicting Justice *B*	dicting Justice *B*
Beverly McLachlin	Claire L'Heureux-Dubé	259/302	0.8576	151/188	0.8032
Frank Iacobucci	John Sopinka	93/123	0.7561	107/142	0.7535

Table 8 produces a complete list of pairwise linear classifications from the Lamer natural Court of 1992-7. Overall, the model does a good job of predicting authorship. Of the seventy-two possible pairings, the model achieved an accuracy rate of at least 70% in fifty-nine pairings (representing 82% of pairings). The lowest pairwise accuracy rate was 0.60, comparing Justice Gonthier (Justice A) with Justice L'Heureux-Dubé (Justice B). Conversely, the highest pairwise comparison accuracy rate was 94%, comparing Justice L'Heureux-Dubé (Justice A) with Justice Cory. As the table illustrates, the prediction rates were higher for some justices than for others. For example, the accuracy of prediction for Justice McLaughlin (Justice A) was at least 85% for each of the eight pairwise comparisons; conversely, the accuracy for Justice Major (Justice A) never exceeded 82% and was as low as 55% (pairwise comparison with Justice McLachlin).

[Note: The following table/form is too wide to be printed on a single page. For meaningful review of its contents the table must be assembled with part numbers in ascending order from left to right. Row numbers, which are not part of the original data, have been added in the margins and can be used to align rows across the parts.]

****
************** This is piece: 1
****

1	Table 8: Linear classifier, sitting justices during Lamer natural Court (1992-7)								
2		V4 Score	Antonio Lamer	Gerard La Forest	Claire L'HeureuxDubé	John Sopinka	Charles Gonthier	Peter Cory	Beverley McLachlin	Frank Iacobucci
3	Antonio Lamer	4.32	_____	0.0764	0.3032	0.0007	0.9269	0.3727	0.6835	0.9457
4	Gérard La	3.77	0.9236	_____	0.8255	0.0270	0.9998	0.8871	0.9718	0.9998

		V4 Score								
	Forest									
5	Claire L'Heureux-Dubé	4.07	0.6968	0.1745	_____	0.0009	0.9916	0.6124	0.8648	0.9956
6	John Sopinka	3.17	0.9993	0.9730	0.9991	_____	1.0000	0.9997	1.0000	1.0000
7	Charles Gonthier	4.94	0.0731	0.0002	0.0084	0.0000	_____	0.0188	0.2316	0.6013
8	Peter Cory	4.20	0.6273	0.1129	0.3876	0.0003	0.9812	_____	0.7738	0.9865
9	Beverley McLachlin	4.56	0.3165	0.0282	0.1352	0.0000	0.7684	0.2262	_____	0.8026
10	Frank Iacobucci	5.02	0.0543	0.0002	0.0044	0.0000	0.3987	0.0135	0.1974	_____
11	John C. Major	4.33	0.5196	0.2478	0.2478	0.0000	0.9695	0.3711	0.7148	0.9844
12	Note: V4 Score column represents V4 Score for each justice. Remaining cells represent P((Row Justice) < (Column Justice)).									

****
*************** This is piece: 2
****

1	
2	John C. Major
3	0.4804
4	0.956

	6
5	0.7522
6	1.0000
7	0.0305
8	0.6289
9	0.2852
10	0.0156
11	___
12	

   It is also important to note that the interpretation of the linear classifier differs slightly from that of the variability scores. Justices with variability scores similar -- or even identical -- to one another may, nevertheless, have writing styles clearly distinguishable from one another. For example, Chief Justice Lamer and Justice Major have nearly identical variability scores of 4.32 and 4.33, respectively. The linear classifier, however, predicted with 88% accuracy the authorship of Chief Justice Lamer's opinions and 75% **185** accuracy the authorship of Justice Major in a pairwise comparison. Similar (or identical) variability scores are consistent with significant differences in how justices use the various function words. For example, while these two justices use the word 'an' at nearly identical rates (mean = 0.61%), Chief Justice Major uses the word 'its' (mean = 0.22%) approximately 1.4 times as often as does Justice Major.

   The purpose of the linear classifier is to show that the text of justices' opinions are often statistically distinguishable from one another, even when the variability scores between two justices do not significantly differ from one another. Of course, identifying the author of a Supreme Court opinion is, in most instances, an academic exercise, given that most majority opinions identify the authoring justice. The linear classifier does, however, have potential applications where the authorship is unknown. *Per curiam* opinions, for example, identify the justices in the majority coalition, but not the author. Similarly, co-authored opinions identify a smaller subset of justices who crafted the opinion, but not the allocation of writing among the justices.

V *Discussion*

This article had both a statistical and an institutional objective. The statistical objective was to construct a general approach to evaluating judicial opinions of the Supreme Court of Canada. By relying on common function words rather than case-specific or law-specific terms, we produced measures of the justices' writing variability that allowed us to make meaningful comparisons of justices, both within and across individual justices as well as cohorts. Comparing within justices, over 75% of justices displayed writing variability that was statistically more variable in their final five years on the Court than in their first five years. And in 83% of pairwise comparisons of the last Lamer natural court, our linear classifier model identified the correct author at least 70% of the time.

Our second, and arguably more significant, objective was to draw inferences from the textual analysis to better understand how justices on the Court produce written opinions. Figure l [FN41] illustrates that prior to the implementation of judicial clerks, the justices' variability scores were lower, consistent with the view that justices performed more of their own writing. The increased variability -- indicated by the upward increased slope in the regression line -- is consistent with the view that recent justices rely more on their clerks when writing opinions. It bears noting, **186** however, that the upward slope in variability scores in Figure l [FN42] over time, prior to the implementation of clerks, suggests an additional factor influencing the use of function words. The most likely explanation is that justices during this period may have increased the practice of quoting other sources -- e.g., dicta in other opinions; statutory provisions that result in increased variability scores.

Also consistent with this point, Figure 2 shows that many justices significantly vary in their variability scores from year to year, [FN43] reflecting sensitivity to the turnover of clerks. Our variability bootstrap scores provide point estimates of the changes in the variability scores over time. [FN44] Chief justices are disproportionately represented among justices who had a higher variability score for the final five years of their tenure than for their first five years, which suggests that the administrative demands of being the chief justice make it more difficult for these justices to focus on opinion writing, meaning that their clerks take on a greater role.

Our earlier point that our textual analysis provides only circumstantial evidence of clerks' actively participating in the opinion writing stage bears repeating. Our central premise throughout the article is that the lower a justice's variability score, the less her or his reliance on their law clerks. We readily acknowledge that, even justices who write their own judgments might vary in their variability of writing style. It is also possible that the differences we do observe in variability scores reflect differences not in the behaviour of the justices but in the ability of the clerks to mimic their justices' writing style. These competing explanations, while conceivable, are undermined when looking at the period before the implementation of clerks, [FN45] the year in which Justice Judson wrote opinions without the assistance of a clerk, [FN46] and the lower variability of contemporary judges who are widely known to write their opinions without the assistance of their clerks. [FN47]

In our main analysis, we focus on majority opinions rather than concurring or dissenting opinions, for largely technical reasons [FN48] due to the fact that concurring and dissenting opinions tend to be shorter than majority opinions. In Figure 4, however, we graph variability scores for dissenting and concurring opinions. The close correlation in variability scores, by justice, between concurring and dissenting opinions lends **187** support for the proposition that justices take a more active role when writing these shorter opinions. [FN49] They may do so because the advantages of delegation are smaller; that is, they could write this opinion in the same time that it would take to delegate it to the clerk and subsequently review it -- or consumptive value of writing such an opinion is higher. [FN50]

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

Figure 4: Variability Scores, concurring and dissenting opinions by individual justice

Our findings on the linear classifier found that the justices' writing styles were sufficiently distinct to allow us, in the vast majority (82%) of pairwise matchings, to correctly identify authorship. It is worth noting ***188** that this rate of accuracy, while high, was lower than for our analysis of justices on the US Supreme Court, which had an accuracy rate of 94%. [FN51] The most plausible explanation for this difference is that the US Supreme Court justices each have *four* clerks, compared with *three* for Supreme Court of Canada justices. The more clerks involved in opinion-writing, all things being equal, the greater the variability in writing; and accordingly, the more distinguishable their writing styles.

More generally, Supreme Court of Canada justices follow trends similar to the US Supreme Court justices: greater variability in writing with the inception of law clerks and a general upward trend among current and recent justices. At the same time, the variability scores for US justices are, on the whole, lower than for the Canadian justices. At first blush, this finding seems at odds, given that each US justice has more clerks than does each Canadian justice. While there are certainly other dimensions in which the two courts are distinguishable -- e.g., judicial ideology [FN52] -- our focus on common function words should be orthogonal to ideology. Canadian justices, while exhibiting a more variable writing style individually, appear more similar to one another along these function words. The explanation, if true, goes beyond this article but merits further consideration.

Much of the criticism targeted at the US Supreme Court is inapplicable to the Canadian Supreme Court. For example, Canadian justices are required to retire at age 75, [FN53] whereas US justices can serve as long as they like. [FN54] Canadian justices may hear cases in panels smaller than nine, allowing them to hear more cases if needed; US justices always sit collectively, limiting their ability to manage their docket. The Court, however, does not appear constrained by its docket, as its caseload has been steadily declining in recent years. [FN55]

Nevertheless, the common institutional characteristic of both Courts is a bimodal workforce with no middle: senior jurists, typically at least fifty years old, coupled with recent law graduates, typically under thirty.

*189 VI *Conclusion*

This article engaged in statistical analysis of judicial opinions in an effort to better understand the Supreme Court of Canada. Our findings provide empirical support for anecdotal accounts that recent and current justices rely more on law clerks in writing opinions than did their predecessors. The normative implications of this finding extend beyond the scope of this article, but it is a discussion worth having, given the importance that judges, practitioners, and legal academics attach to the written words of each opinion.

More broadly, this article seeks to build interdisciplinary approaches to understanding jurisprudence and other areas of traditional legal scholarship. The advances in statistical computation and analysis make it possible to explore systematically questions in jurisprudence and doctrine in a way that can inform how we think both descriptively and normatively about the law.

*190 Appendix

Table Al: Current and departed justices, pre-1949

Num	Justice	Province	Appointing Prime Minister	Year Joined Court	Year Left Court	Years On Court	Opinions	Awage Word Length	Vari-

									abi lity Sc ore
39	Charles Holland Locke	BC	King	1947	1962	15	196	2864	3. 32
40	James Wilfred Es- tey	SK	King	1944	1956	12	214	4331	4. 15
41	Roy Lindsay Kellock	ON	King	1944	1958	14	158	2290	2. 86
42	Ivan Cleveland Rand	NB	King	1943	1959	16	240	1757	2. 45
43	Albert Blellock Hudson	MB	King	1936	1947	11	18	1227	2. 49
44	Patrick Kerwin	ON	Bennett	1935	1963	28	293	1650	2. 65
45	Henry Hague Dav- is	ON	Bennett	1935	1944	9	41	2412	2. 64
46	Frank Joseph Hughes	ON	Bennett	1933	1935	2	7	2399	2. 41
47	Oswald Smith Crocket	NB	Bennett	1932	1943	11	21	2837	3. 33
48	Lawren ce Arthur	QC	King	1930	1939	9	7	2003	2.

	Dumoulin Cannon							69
49	Robert Smith	ON	King	1927	1933	6	19	1814 2.19
50	John Henderson Lamont	SK	King	1927	1936	9	24	2485 3.36
51	Edmund Leslie New-combe	NS	King	1924	1931	7	22	2378 2.65
52	Pierre-Basile Mig-nault	QC	Borden	1918	1929	11	30	1583 2.21
53	Louis-Philippe Brodeur	QC	Laurier	1911	1923	12	32	875 2.05
54	Francis Alexander Anglin	ON	Laurier	1909	1933	24	128	1869 2.41
55	Lyman Poore Duff	BC	Laurier	1906	1944	38	306	4260 3.64
56	Sir Charles Fitzpatrick	QC	Laurier	1906	1918	12	74	1304 2.1
57	James Maclennan	ON	Laurier	1905	1909	4	10	2093 2.5
58	John Id-ington	ON	Laurier	1905	1927	22	109	1856 2.29

59	Albert Clements Killam	MB	Laurier	1903	1905	2	10	1384	2. 03
60	Wallace Nesbitt	ON	Laurier	1903	1905	2	8	2332	2. 28
61	John Douglas Armour	ON	Laurier	1902	1903	1	2	574	1. 13
62	David Mills	ON	Laurier	1902	1903	1	9	1095	2. 05
63	Sir Louis Henry Davies	PE	Laurier	1901	1924	23	108	1 772	2. 34
64	Désiré Girouard	QC	Bowell	1895	1911	16	17	1 310	1. 89
65	George Edwin King	NB	Thompson	1893	1901	8	10	1 352	2
66	Robert Sedgewick	NS	Thompson	1893	1906	13	25	1 998	2. 31
67	Christopher Salmon Patterson	ON	Macdonald	1888	1893	5	8	1 558	2. 23
68	John Wellington Gwynne	ON	Macdonald	1879	1902	23	83	3 641	4. 34
69	Sir Henri-	QC	Mackenzie	1878	1906	28	43	1 624	2.

	Elzéar Taschereau								74
70	William Alexander Henry	NS	Macken zie	1875	1888	13	46	2 634	2. 73
71	Jean-Th omas Taschereau	QC	Macken zie	1875	1878	3	2	2 505	5. 19
72	Samuel Henry Strong	ON	Macken zie	1875	1902	27	150	2 163	2. 73
73	William Johnstone Ritchie	NB	Macken zie	1875	1892	17	87	1 963	3. 02
74	Sir Wil- liam Buell Richards	ON	Macken zie	1875	1879	4	7	5 075	3. 25

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

*192 Figure A1: Variability scores by original language of opinion

Note: The blue curve is a graph with mean zero and variance 1. The red curve is a graph with mean zero and variance 1.1. Most of the probability distribution falls between 1 and-1, where the blue curve is larger. In this example, approximately 70% of the points (opinions) chosen from either distribution will be classified as belonging to the blue curve. Accordingly, points from the blue distribution will be correctly classified about 70% of the time, while points from the red distribution will be correctly classified about 30% of the time.

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

Figure A2: Shape and overlap of the distribution between two justices

[FN1]. David Margolick, 'At the Bar: Annual Race for Clerks Becomes a Mad Dash with Judicial Decorum Left in the

Dust,' *The New York Times* (17 March 1989) B4.

[FN2]. Information provided by the Supreme Court of Canada, *About the Court* (October 2012), online: <http://www.scc-csc.gc.ca/court-cour/administration/index-eng.asp>; see also Lome Sossin, 'The Sounds of Silence: Law Clerks, Policy Making and the Supreme Court of Canada' (1996) 30 UBC L Rev 279 at 283 [Sossin], describing the number of clerks in each chambers.

[FN3]. See FL Morton & Rainer Knopff, *The Charter Revolution and the Court Party* (Toronto: University of Toronto Press, 2000) at 110 [Morton & Knopff], describing the burgeoning power of law clerks at the Supreme Court of Canada.

[FN4]. For example, McInnes et al report that over 75% of clerks during the early 1990s attended one of seven law schools (Toronto, McGill, Osgoode Hall, Ottawa, Dalhousie, Alberta, and the University of British Columbia); Mitchell Mclnnes, Janet Bolton, & Natalie Derzko, 'Clerking at the Supreme Court of Canada' (1994) 33 Alta L Rev 58 [Mclnnes et al].

[FN5]. See e.g. Donald R Songer, *The Transformation of the Supreme Court of Canada: An Empirical Examination* (Toronto: University of Toronto Press, 2008).

[FN6]. See e.g. *Transcriptions of Conversations between Justice William O Douglas and Professor Walter F Murphy* (Cassette No 3, 20 December 1961) (Princeton, NJ, Princeton University Library, 1981), where Justice Douglas states, 'I have written all my own opinions'; see also Edward Lazarus, *Closed Chambers* (New York: Crown, 1998) at 271, identifying Justices Scalia and Stevens as writing their own opinions.

[FN7]. See Todd C Peppers, *Courtiers of the Marble Palace: The Rise and Influence of the Supreme Court Law Clerk* (Stanford, CA: Stanford University Press, 2006) at 58, 96-7, 114 [Peppers], describing the relationships of Justices Holmes, Cardozo, and Douglas to their clerks.

[FN8]. See Mark Tushnet, 'Thurgood Marshall and the Brethren' (1992) 80 Geo LJ 2109 at 2112, describing how Marshall relied more heavily on his law clerks early in his tenure.

[FN9]. See Jeffery S Rosenthal & Albert H Yoon, 'Judicial Ghostwriting: Authorship on the US Supreme Court' (2011) 96 Cornell L Rev 1307 at 1323-4 [Rosenthal & Yoon, 'Judicial'], showing, in Figure 2, the year-to-year variability scores of these justices among others.

[FN10]. See Morton & Knopff, supra note 3 at 110, describing the burgeoning power of law clerks at the Supreme Court of Canada.

[FN11]. See ibid at 111, describing how a law clerk for Chief Justice Dickson purportedly created part 1 of the Oakes Test.

[FN12]. See McInnes et al, supra note 4 at 78: 'any suggestion that Canada's Justices abdicate their responsibilities for writing judgments is false; in every instance, a decision of the Court ultimately and fundamentally is the product of the nine men and women who sit on the bench.'

[FN13]. See Sossin, supra note 2 at 297-8.

[FN14]. See Peter McCormick & Ian Greene, *Judges and Judging: Inside the Canadian Judicial System* (Davidson, NC: Lorimer Press, 1990) at 208. This viewpoint is echoed by-Justice Bertha Wilson; see Bertha Wilson, 'Decision-Making

in the Supreme Court' (1986) 36 UTLJ 227 at 236, describing how clerks' central role is to provide background research.

[FN15]. See e.g. Axel Börsch-Supan & Matthias Weiss, 'Evidence from Work Teams at the Assembly Line' (2011) [Unpublished working paper], online: < http://www.mea.mpisoc.mpg.de/uploads/user_mea_discussionpapers/1057_MEA-DP_148-2007.pdf>, finding that productivity among auto workers does not decline as they grow older, which suggests that any physical decline may be offset by experience and an ability to collaborate with others; Ray C Fair, 'How Fast Do Old Men Slow Down?' (1994) 76 Review of Economics and Statistics 103-8, describing physical decline among professional athletes.

[FN16]. See Rosenthal & Yoon, Judicial,' supra note 9 at 1313-7.

[FN17]. Scholars have applied **stylometry** to literature, see e.g. O Seletsky et al, 'The Shakespeare Authorship Question' (2007) [unpublished, Dartmouth College]; political texts, see e.g. Frederick Mosteller & David L Wallace, *Inference and Disputed Authorship: The Federalist* (Boston, MA: Addison-Wesley, 1964) [Mosteller & Wallace]; and they have applied it to speeches, see e.g. Edoardo M Airoldi et al, 'Who Wrote Ronald Reagan's Radio Addresses?' (2006) 1 Bayesian Analysis 289.

[FN18]. Other approaches of writing style -- e.g. sentence length, paragraph length, punctuation -- yielded comparable results.

[FN19]. David Madigan et al, 'Author Identification on the Large Scale' (paper delivered at the Classification Society of North America, 2005) [unpublished, on file with authors].

[FN20]. See Mosteller & Wallace, supra note 17.

[FN21]. The seven words - 'every,' 'my,' 'shall,' 'should,' 'upon,' 'will,' 'you' -- were each less than 0.1% of all words in the Court's majority opinions.

[FN22]. See, e.g., Peppers, supra note 7 at 18-20, describing clerks' reluctance to discuss their relationship with their justices for reasons of confidentiality.

[FN23]. Judge Posner's variability measure was 2.60 for the period 1981-2010, with a standard deviation of 0.14, while Judge Easterbrook's was 2.42, with a standard deviation of 0.18; see Rosenthal & Yoon, Judicial,' supra note 9 at 1325. Their variability scores were lower than for any contemporary justices (for those on the Court as of 2010) on the US Supreme Court, which ranged from 3.06 (Breyer, J) to 3.73 (Kennedy, J); their standard deviations were similarly lower than those of the other justices, which ranged from 0.56 (Stevens, J) to 0.62 (Kennedy, J); see ibid at 1323.

[FN24]. Lexum, *Judgments of the Supreme Court of Canada*, online: <http://scc.lexum.org/decisia-scc-csc/scc-csc/en/nav.do>.

[FN25]. For justices who joined the Court prior to 1949, see Appendix, Table A1.

[FN26]. For the software used to download and analyse the text and a description of the software program, see Jeffrey S Rosenthal, 'Explanation of the Software at: http://probability.ca/scc/' (2012), online: <http:// probability.ca/scc/README>.

[FN27]. See Carissima Mathen, 'Dissent and Judicial Authority in Charter Cases' (2003) 52 UNBLJ 321 at 323, noting that the Court 'rarely issues *per curiam* decisions.' Our analysis of the data reveals that co-authored opinions similarly

occur infrequently.

[FN28]. For a comparison of variability scores by the justices' original language of opinion, see Figure A2 in the Appendix.

[FN29]. Both quoted rates are significantly distinct from 0, and from each other, at $p < 0.001$.

[FN30]. See Songer, supra note 5 at 139.

[FN31]. Justice Judson did not have a law clerk during the 1974-5 term. See Michael J Herman, 'Law Clerking at the Supreme Court of Canada' (1975) 13 Osgoode Hall LJ 279, reporting this fact.

[FN32]. See Bradley Efron & Robert J Tibshirani, *An Introduction to the Bootstrap* (London: Chapman and Hall, CRC, 1994), describing the bootstrapping approach.

[FN33]. We include replacement in accordance with convention -- see ibid; in addition, because of the relatively low number of opinions written by some justices, sampling without replacement would cause some justices to drop from our analysis.

[FN34]. Each iteration of the bootstrap score produces a slightly different variability score, but each is similar to the original variability score. For this reason, we elect to report the original variability score.

[FN35]. See generally Lee Epstein et al, 'Ideological Drift among Supreme Court Justices: Who, When, and How Important?' (2007) 101 Nw UL Rev 1383.

[FN36]. See David J Garrow, 'Mental Decrepitude on the US Supreme Court: The Historical Case for a 28th Amendment' (2002) 67 U Chicago L Rev 995 at 995.

[FN37]. See *Constitution Act*, 1867 (UK), 30 & 31 Vict, c 3, s 91, reprinted in RSC 1985, App II, No 5, s 99 (2).

[FN38]. See Rosenthal & Yoon, 'Judicial,' supra note 9 at 1332-3; we posed this question to US Constitutional scholars.

[FN39]. See Jeffrey S Rosenthal & Albert H Yoon, 'Detecting Multiple Authorship of United States Supreme Court Legal Decisions Using Function Words' 5 Annals of Applied Statistics 283 at 287, describing how over-fitting the data is where one constructs a model that fits the existing data well but cannot effectively predict using new data.

[FN40]. For an illustration of how the shape and overlap of the distribution between two justices determines the degree of symmetry, see Figure A2 in the Appendix.

[FN41]. See Figure 1 above.

[FN42]. Ibid.

[FN43]. Ibid.

[FN44]. See Table 5 above.

[FN45]. See Figure 1 above.

[FN46]. See Figure 3 above.

[FN47]. See Rosenthal & Yoon, Judicial,' supra note 9 at 1325-6.

[FN48]. See text accompanying note 26.

[FN49]. This result is consistent with the findings for concurrences and dissents for the US Supreme Court; see Rosenthal & Yoon, 'Judicial,' supra note 9 at 289.

[FN50]. See e.g. Justice Thurgood Marshall, explaining judicial dissents: 'I enjoy the fight. I agree with the old saying, "I love peace but I adore a riot." You've got to be angry to write a dissent'; qtd in Leon Friedman & Fred L Israel, *The Justices of the United States Supreme Court 1789-1978* (New York: Chelsea, 1980), vol 5 at 50 *[sub verba* 'Thurgood Marshall'].

[FN51]. See Rosenthal & Yoon, 'Judicial,' supra note 9 at 1336.

[FN52]. See Benjamin Alarie & Andrew Green, 'Should They All Just Get Along: Judicial Ideology - Collegiality, and Appointments to the Supreme Court of Canada' (2007) 58 UNBLJ 73 at 84: 'Given the "brokerage" model of politics in Canada in the past and the lack of significant differences in policy preferences in most areas across parties (particularly in the 1980s and 1990s), the appointees to the Court may have been largely similar ideologically.'

[FN53]. See supra note 37 and accompanying text.

[FN54]. See US Const, art III § 1.

[FN55]. See Peter McCormick, 'Standing Apart: Separate Concurrence and the Modern Supreme Court of Canada, 1984-2006' (2008) 53 McGill LJ 137 at 166, describing recent caseload trends on the Court.

63 U. Toronto L.J. 159

END OF DOCUMENT