Aesthetic Preferences and Policy Preferences as Determinants of U.S. Supreme Court Writing Style[*]

Jeffrey Budziak[†]   Daniel Lempert[‡]

June 29, 2021

Abstract

Recent literature on writing style in U.S. Supreme Court opinions has focused on style as a means of furthering justices' policy goals. In particular, an opinion's clarity is proposed to make the implementation of the announced policy more likely. We give a formal argument that the observed distribution of opinion clarity is not easily reconcilable with justices who are striving to write clearly in service of policy implementation related goals; this is true even if there are case-level costs that sometimes make writing clearly more difficult. We propose that justices having aesthetic preferences—essentially, stylistic preferences over opinion language that are unrelated to policy implementation—that they weight heavily could explain the observed distribution of opinion clarity. Our analysis of some 4500 majority opinions 1955–2008 is largely consistent with our theoretical argument.

[†]Associate Professor of Political Science, Western Kentucky University. email: jeffrey.budziak@wku.edu.

[‡](corresponding author) Associate Professor of Politics, SUNY Potsdam. email: lemperds@potsdam.edu.

A lively debate in judicial politics concerns the extent to which judges seek to enact their policy preferences in the face of competing goals. Foundational work on the U.S. Supreme Court viewed the fundamental, perhaps exclusive, goal of justices to be the enactment of their policy preferences (for one overview, see Baum 1997, Ch.2). Subsequent work called attention to the constraints justices face as they seek to enshrine their preferences over legal policy into law (Epstein and Knight 1998; Maltzman, Spriggs and Wahlbeck 2000). Even more recently, Baum (2006) emphasized that goals other than the implementation of policy preferences—in fact, goals that are not even instrumental to the implementation of justices' preferred policies—are likely important to justices (see also Posner 1993).

Here, we are interested in justices' goals and constraints as they choose the writing style of their opinions. Perhaps the most completely theorized work on this topic, Black, Owens, Wedeking and Wohlfarth (2016*b*), emphasizes that writing style can advance the goal of policy preference implementation. In particular, the argument is that clearly-written opinions help majority opinion authors implement their policy preferences by inducing compliance with the opinion and by managing the Court's legitimacy (Black et al. 2016*b*, 17-39). We refer to this as the "conventional theory" (of opinion clarity). Our aim is not to directly reassess the empirical results backing the conventional theory. Instead, we propose that more attention should be given to constraints that could prevent justices from writing clearly, and goals besides policy preference implementation that justices may seek to achieve through their writing style. Particularly, we suggest that justices' *aesthetic preferences*—their views about what constitutes good writing style, independent of concerns about policy implementation— may explain a substantial amount of variation in opinion clarity.

Theory

We motivate our argument with a simple empirical observation, displayed in Figure 1: the clarity of Supreme Court majority opinions varies considerably. In Figure 1, and in our subsequent analyses, we follow recent scholarship that defines a clearly-written opinion as

one that is easily readable (e.g., Black, Owens, Wedeking and Wohlfarth 2016*a*; Black et al. 2016*b*; Owens, Wedeking and Wohlfarth 2013). The particular metric we use is the Flesch-Kincaid Grade Level (FKGL), a widely-utilized measure of readability based a text's sentence and word length; conveniently, it is scaled to approximate the (U.S.) grade level of education required to understand the text.[1]
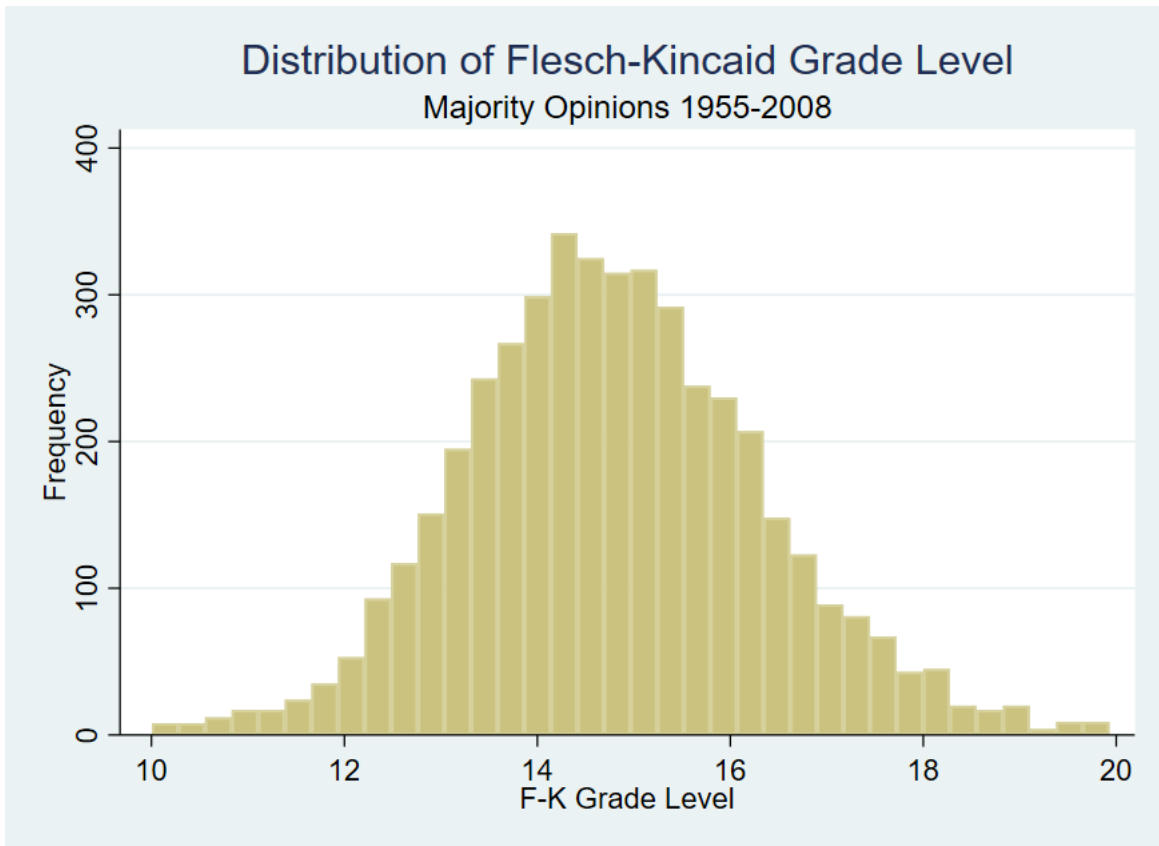

Figure 1. Number of Supreme Court opinions by level of clarity.

This empirical distribution is surprising under the most straightforward interpretation of the conventional theory of opinion clarity. If opinion clarity matters insofar as it aids implementation of the policy announced in the opinion, the author's utility u from any given opinion o written at level of clarity w is

$$u_o = w.$$

---

[1]We discuss the measure, and the sample of opinions in Figure 1 in more detail below.

The utility-maximizing choice obviously is then to always write at the maximum level of clarity. But this does not come close to describing reality: as Figure 1 shows there is substantial variation in majority opinion clarity; moreover, clarity is about normally distributed, i.e., the modal opinion is about as close to the clearest-possible SC opinion as it is the least-clear possible opinion.

As Black et al. (2016b) recognizes in its empirical analyses, a case's attributes may also affect opinion clarity. In the terms of our formal analysis, we would say that that costs and benefits associated with a case's attributes may also affect a justice's utility from writing a given opinion at given level of clarity. For example, justices might derive relatively more benefit from writing a salient case clearly, if they value compliance in salient cases more than compliance in nonsalient cases. Or, it may be that it is more difficult, and thus costly, to write a case clearly as the end of the term approaches and time pressures increase. Can this explain the observed distribution of opinion clarity?

To write down the author's utility function, let c denote the n case attributes that affect a justice's utility from writing at level of clarity w. Let $v_1, \ldots, v_n$ denote the functions that assign utility from writing an opinion with clarity w for the each of the n case attributes observed in case o.[2] Then, symbolically,

$$u_o = w + v_1(c_{(o,1)}, w) + \ldots + v_n(c_{(o,n)}, w).$$

Note that if—as commonly assumed in empirical analyses—the v_i are linear in w, whether increasing or decreasing, the optimal w is either the maximum or minimum level of clarity. (This is because any sum of linear functions is itself linear, implying that u is linear in w.) But this is not concordant with the empirically-observed distribution of opinion clarity, which is approximately normal. Thus, a straightforward addition of case-based costs and benefits into justices' utility functions is not enough, by itself, to yield the observed distribution of

---

[2]As above, we leave in w in the utility function; this can be understood as the baseline benefits, from compliance and legitimacy, of writing an opinion with clarity w.

opinion clarity.

Alternatively, if all v_i are all monotonically increasing in w, the optimal choice of w is to write at the maximum level of clarity. (This is because the sum of any monotonically increasing functions is also monotonically increasing, which implies u is monotonically increasing in w.) Of course, this is not reflected in the empirical distribution of opinion clarity either.

Finally, even if individual v_i are not linear or monotonically increasing, if u_o is itself monotonic in w, the optimal choice is again either to write at the maximum or minimum level of clarity. This again appears not be the case, empirically. So, costs and benefits entering into justices' utility functions, in any of these forms, are not enough to explain the observed distribution of opinion clarity.

How can we reconcile the observed distribution of opinion clarity with these conclusions? One obvious way that u can be nonmonotonic in w is if it includes a term—especially a relatively influential one—that is itself nonmonotonic in w. Consider the possibility that justices have preferences about the level of clarity an opinion should have, independent of case-specific costs or concerns about implementation. We refer to such preferences over writing style as *aesthetic*.

Fundamental to our conceptualization of aesthetic preferences is that justices have preferences over writing style that are distinct from the pursuit of policy goals and the tradeoffs frequently made to achieve those goals. Most centrally, aesthetic preferences reflect a judge's ideal vision of style—the style of opinion writing a justice finds most intrinsically pleasing, absent all other considerations.

There are strong reasons to believe that justices have such preferences. All justices are thoroughly socialized into the legal profession, which puts strong emphasis on writing. Most have shared experiences as either law clerks or judges on other courts (or both) where much of their time and effort was dedicated to writing opinions. Given the centrality of legal writing to justices' pre-Court careers, it would be surprising if they did not have relatively strong preferences over what constitutes good writing.

We can cite qualitative evidence that U.S. appellate judges do in fact care about writing style, and not just as a means of policy implementation. Articles written by federal judges about opinion writing are too numerous to list here, but Vance's (2011) annotated bibliography compiles and describes numerous such pieces. The Federal Judicial Center's (2013) official guide to opinion writing for judges is written primarily by federal judges and includes one chapter devoted entirely to questions of style. In 2006 and 2007, eight of the nine sitting justices sat for substantial interviews about legal writing style with the linguist Bryan Garner (2011). And almost a century earlier, then-judge Benjamin Cardozo defended his views on judicial writing style at length, arguing that opinions can be seen as a form of literature (Cardozo 1931).[3] Notably, he declined to endorse clarity as the sole marker of strong judicial writing, writing that "clearness, though the sovereign quality, is not the only one to be pursued (Cardozo 1931, 6,9)."

We can expand our conception of aesthetic preferences by considering that judicial opinions are ultimately written to be read by others. Baum (2006, 28) theorizes that judging can be understood as an act of self-presentation. Of all the activities that encompass the act of judging, opinion writing is perhaps the easiest to understand from this perspective: opinions are the ultimate product on which judges are evaluated by audiences important to them. While self-presentation can take several forms, most relevant for our conceptualization of aesthetic preferences is what Baum calls "personal self-presentation." People "engage in self-presentation because they seek popularity and respect as ends in themselves, not as means to other ends (Baum 2006, 29)." In the case of writing style, justices may deviate from their ideal, inherently most-preferred writing style to write in ways that they expect groups important to their social identity will prefer. We subsume such considerations under the umbrella of "aesthetic preferences" as well. The most fundamental point about aesthetic preferences is that they are not shaped by considerations directly related to policy implementation.

---

[3]Cardozo's article "Law and Literature" was originally published in the July 1925 issue of *The Yale Review*. The citation here is to a book, easier to locate online, that reprints the essay.

There are good reasons to expect that aesthetic preferences over writing style will differ across justices. True, Supreme Court justices have much in common. All are highly educated, most at elite private institutions. All, by virtue of their appointment to the Court—and most, well before it—are members of an elite political class. All have been trained and socialized in the legal profession and most of the justices in our sample have similar career paths. Still, common educational levels, professional training, and career background do not guarantee uniformity of thought. After all, justices vary widely in their preferences over legal policy even though their backgrounds are similar in all these respects. The same is likely true for aesthetic preferences over writing style: even among similarly well-educated people, among people sharing backgrounds and occupations, both Hemingway (3.30) and Hawthorne (15.90) have their admirers.[4] Indeed, Varsava (2021) argues that judges emphasize writing in an "aesthetically pleasing" and "distinct, personal" style to such a degree that the integrity of the judicial role and the legitimacy of opinions can be undermined.

Qualitative and journalistic accounts also indicate that justices' aesthetic preferences vary. For example, Justice Douglas' opinions have been described as "conversational, like a person talking intimately to another person," and one clerk recalls a laborer stating that Douglas' opinions are the only ones that he can understand (Small 2007). In contrast, Justice Souter's style has been called "turgid" and so "ornate" and "convoluted" that it caused the justice "difficulty in communicating his ideas"; there is even indication that Souter understood that his style was perceived this way and nonetheless preferred it (Rosen 1993; Rosen 2009). Moreover, certain quantitative text analyses assessing individual judges' writing style, as measured by the relative use of function words, are also consistent with the proposition that aesthetic preferences vary (Carlson, Livermore and Rockmore 2016; Frankenreiter 2019; Rosenthal and Yoon 2011).[5]

---

[4]According to Dodson and Dodson (2015), both authors have been cited in Supreme Court opinions. These FKGL scores are for excerpts from *The Old Man and the Sea* and *The Scarlet Letter*, as presented in Dalvean and Enkhbayar (2018), which also gives scores for dozens of other works of literature.

[5]Function words are words that have little substantive content, for example, *a*, *but*, and

Judges themselves recognize that preferences over style vary. Posner (1995, 1426) states that "anyone who has read a large number of judicial opinions [...] will have noticed that judicial style is not uniform." He then goes on to distinguish between opinions that tend to be "lofty, formal, imperious, impersonal, 'refined,' ostentatiously 'correct' [...] even hieratic" and those that are "direct, forthright, [...] colloquial, informal, [...] even demotic." Cardozo's (1931, 10) taxonomy is more elaborate:

> I seem to discern six types or methods which divide themselves from one another with measurable distinctness. There is the type magisterial or imperative; the type laconic or sententious; the type conversational or homely; the type refined or artificial, smelling of the lamp, verging at times upon preciosity or euphuism; the type demonstrative or persuasive; and finally the type tonsorial or agglutinative, so called from the shears and the pastepot which are its implements and emblem.[6]

The Federal Judicial Center (2020, 84) is more prosaic in its instructions to new clerks: "Each judge has a different writing style. Some prefer simple declarative sentences and use plain language. Others employ complex sentences and a varied vocabulary."[7]

Nor are questions of writing style purely abstract for judges. The rather indecorous exchange between circuit judges Posner (1995, 1437–1443) and Wald (1995) shows that judges can be quite sensitive when challenged about the stylistic choices they make in their opinions. In sum, we can cite at least anecdotal evidence that judges' aesthetic preferences over writing style vary, and that they feel strongly about these preferences.

Formally, we can take aesthetic preferences into account with the following modified utility

---

the (see Rosenthal and Yoon 2011, 287).

[6]Granted, it is probably not feasible, at scale and in practice, to classify opinions into these categories (even once one looks up *tonsorial* and *agglutinative*). But the relevant point is that opinions across these categories are sure to vary across the single dimension of clarity, as political science scholarship has conceptualized it.

[7]Notably, the guide goes on to recommend that clerks adopt the writing style their judge prefers when drafting or editing opinions.

function. Denoting justice j's aesthetically most-preferred level of clarity as w_j^*, and letting k be a nonnegative weight,

$$u_{(j,o)} = w + v_1(c_{(o,1)}, w) + \ldots + v_n(c_{(o,n)}, w) - k(w_j^* - w)^2. \tag{1}$$

If k is sufficiently large—that is, if justices place sufficient weight on their aesthetic preferences, compared to implementation-related and case-specific considerations—$u_{j,o}$ is non-monotonic in w, which means that the equilibrium level of clarity w varies across cases and justices. As such, aesthetic preferences may well go some ways toward explaining the empirical distribution of opinion clarity we observe. More importantly, our theoretical understanding of justices' stylistic choices would be enriched if aesthetic preferences are shown to be an important factor in justices' writing; we turn to this empirical exercise below.

Of course, we recognize that clarity is but one element of style, and justices are likely to have aesthetic preferences over other elements as well. However, these other elements of style are conceptually independent of clarity as we define it (below), and thus do not enter into the utility function in Eq. 1. For example, justices may vary over their preferences for emotional language (Krewson 2019), but describe similar emotional states as "mad," "angry," "indignant," or "apoplectic." Aesthetic preferences encompass various elements; clarity is an important and oft-discussed one. The strategic deployment of clarity has been widely hypothesized and the concept can be measured at scale. For these reasons, we are here concerned with aesthetic preferences over clarity.[8]

Measurement and Sample

In short, our theoretical approach proposes that justice-specific aesthetic preferences are weighted relatively heavily by justices when compared to policy implementation related goals.

---

[8]This is not to say that studies of other stylistic attributes such as emotional language (Krewson 2019), type-token ratios (Frankenreiter 2019), or function word use (Carlson, Livermore and Rockmore 2016) are uninteresting.

Since there is substantial variance in majority opinion clarity, it is unlikely that the policy-related benefits of clarity override all other costs and benefits in justices' opinion-writing calculus. Still, consistent with the conventional theory of opinion clarity, it may be true that justices seek to write clearly, but are only conditionally motivated by case-specific benefits to do so, or are constrained from doing so by case-specific costs.[9]

If this is so, we should see systematic increases in opinion clarity when the policy implementation related benefits to writing clearly are greater, and when costs are smaller. Below, we delineate a set of case-level covariates that affect the costs and benefits of writing clearly. Later, we will use these covariates to construct two scenarios: one where costs are particularly high and benefits particularly low, and one where benefits are low and costs high. If justice are, in general, attempting to write clearly, but are sometimes prevented from doing so due to case-specific cost-benefit calculations, we should see substantially greater clarity in the low-cost/high-benefit case, compared to the high-cost/low-benefit case.

We have also theorized that the distribution of opinion clarity is plausibly explained by justices' aesthetic preferences, if weighted heavily compared to policy-related costs and benefits. If—after accounting for case-specific considerations—justices still vary considerably in how clearly they write, that would provide evidence that aesthetic preferences play an important role in opinion writing style. Granted, the evidence would not be direct—direct evidence would require an ex ante measure of each justice's most-aesthetically-preferred level of clarity. Still, wide variation across justices, after accounting for case-specific factors, would be concordant with the theoretical proposition that aesthetic preferences are important. And if across-justice variation dominates the impact of case-level costs and benefits, the implication would be that justices are not prevented from writing clearly by contextual constraints (or lack of incentives); rather, their inherent aesthetic preferences over writing style are the key determinants of their outputs.

We turn now to the covariates we include in our analyses to measure case-level costs and

---

[9]Note that could hold even if justices give some weight to their aesthetic preferences.

benefits. We divide the variables into two categories: those that are expected to affect the costs of writing clearly, and those that are expected to affect the benefits to doing so. Unless otherwise noted, the variables are drawn from the Supreme Court Database (Spaeth, Epstein, Segal, Ruger, Martin and Benesh 2017).

Covariates Affecting Costs of Writing Clearly

We first consider some constraints that may make it more difficult for a justice to write clearly. Complex cases may require greater effort to write clearly (e.g., Owens, Wedeking and Wohlfarth 2013, 47). In particular, cases involving multiple legal issues cases may implicate a wider variety of factors (sources of law, precedents to consider, etc.) than cases presenting a single legal issue. As such, it may be more difficult to write clearly when resolving these cases. We thus include the variable *Multiple Legal Issues*, an indicator coded one if multiple issues were identified in the Supreme Court Database, and zero otherwise.

In any given case, both the composition of the majority coalition and its size could affect the clarity of the majority opinion (e.g., Black et al. 2016*b*, 73). When majority coalitions are ideologically diverse, coalition members are more likely to make requests of the majority opinion author (Maltzman, Spriggs and Wahlbeck 2000, 82) and the majority opinion author is likely to take time in an effort to accommodate those requests (Maltzman, Spriggs and Wahlbeck 2000, 116). Similarly, when majority coalitions are small, majority opinion authors are more likely to both receive and accommodate requests (Maltzman, Spriggs and Wahlbeck 2000, 82, 117). Considering and honoring requests from other coalition members may limit the ability of an author to write their opinion as clearly as they might have absent those request. To account for this possibility, we include two variables. The first is *Majority Coalition Diversity*, which equals the standard deviation of the Martin-Quinn (2002) Ideology Scores for all members of the majority coalition. The second is *Number of Majority Votes*, which is the number of votes cast for the majority.[10]

---

[10]We discuss and account for the potential endogeneity of these variables below.

Temporal constraints may also make it more difficult to write clearly. The logic is simple: If justices are devoting their time to other responsibilities, or are otherwise pressed for time, it is relatively more costly to write opinions in the style they prefer. We test this expectation by including several variables. First, *Workload* is the number of other majority opinions a justice is working on concurrently with the majority opinion in question. The more majority opinions a justice has to balance, the less time and effort can be dedicated to any one opinion. Second, we include *Total Docket*, which is the size of the appellate docket on the Court in the term that the opinion was written (including cert denials). We also include *Merits Docket*, which indicates the number of cases decided on the merits in a given term. These measures provide a proxy for the other responsibilities that require a justice's time, including oral argument, opportunities to write separately, etc. Finally, we include a variable, *End of Term*, which captures number of days left until the end of the Supreme Court term.[11] If justices feel pressure to complete their work prior to the end of the term, they may forgo the time and effort necessary to draft an opinion as clearly as they might have, had they more time.

Covariates Affecting Benefits of Writing Clearly

We now turn to covariates that are expected to increase the benefits of writing clearly. Essentially, the argument is that when justices believe a case is particularly important, or otherwise care about policy implementation in a given case, they should be more willing to put forth the effort to write clearly. This logic is consistent with findings in the literature indicating that justices' votes (e.g., Bartels 2011) and their behavior related to bargaining over opinion content (e.g., Epstein and Knight 1998; Maltzman, Spriggs and Wahlbeck 2000) becomes more policy-oriented when they view a case as important. Similarly, a justice might write more clearly in cases they see as important, if—as in the conventional theory—they believe that opinion clarity will make it more likely that the policy announced in the opinion is implemented.

---

[11]Technically, we calculate the distance between (1) the end of the term and (2) the midpoint between the case's oral argument and the opinion announcement.

Scholars generally agree that issues and events political actors feel strongly about (Niemi and Bartels 1985) or view as important (Baird 2004) are best defined in terms of salience. Case salience is an important concept in the judicial politics literature, but the concept is not self-defining. We consider and use three distinct conceptualizations; we seek to measure (or at least proxy) how important a case is to the justices and to audiences that justices care about.

First, we consider the measure proposed in Clark, Lax and Rice (2015). This approach utilizes a latent variable model based on newspaper coverage of a case prior to the Court's decision. For our purposes, a key strength of the metric is that it obviates concerns about endogeneity between the Court opinion's language and the measured importance of the associated case. Clark, Lax and Rice (2015, 40) conceptualizes a case's salience to a Supreme Court justice "as the weight the justice places on the utility she receives from her decision in the case." They argue, and cite to work indicating, that media-based measures, such as theirs, "are an appropriate manifestation of this type of salience (Clark, Lax and Rice 2015, 40-42)." Thus we employ their measure, which we refer to as *Media-Based Salience*, as our first measure of salience.

Next, we consider legal characteristics of cases that make them more salient to justices. Scholars have long argued that cases that are "legally salient" may be particularly important to the justices (e.g., Maltzman, Spriggs and Wahlbeck 2000, 46). The basic claim is that cases that make substantial changes to existing legal frameworks should be relatively more salient. This variant of salience is particularly relevant to opinion writing: If justices expect opinion clarity to help with policy implementation, they should write especially clearly when a policy change occurs, compared to when existing policy remains in place. We use the standard operationalization for the variable *Legal Salience*, coding it so that it equals one if a case overturns precedent or finds a law unconstitutional, and zero if it does neither.

Last, we consider whether justices may get more benefit from writing clearly in cases salient to relevant social and political groups. It is well-established that justices are more likely to vote to hear cases in which organized groups take an interest (Caldeira and Wright 1988).

The theoretical basis for this is that justices see a case as more significant when there is a greater demand for adjudication (i.e., for resolution of the legal dispute). One measure of this demand for adjudication are amicus briefs filed by organized groups; as such, justices can take the presence of amicus briefs as reflecting a case's importance (Caldeira and Wright 1988, 1112). If justices continue to care more, at the merits stage, about cases that are more salient to relevant groups, the benefits of writing clearly in such cases should be relatively greater. Indeed, Maltzman, Spriggs and Wahlbeck (2000, 51, 83, 89, 119, 146) presents several pieces of evidence indicating that justices care more about legal policy in cases with more amicus briefs filed, and that authors put more effort into drafting opinions in such cases. Thus, we include *Group Salience*, based on the number of amicus briefs in case, as reported in Collins (2008) and Box-Steffensmeier and Christenson (2012). Because amicus participation has increased over time, we create a term-specific amicus z-score (e.g., Maltzman, Spriggs and Wahlbeck 2000, 45).

In addition to the above measures of salience, the Court's agenda-setting process gives some clues about how important a given case is to the Court. The majority opinion author frequently indicates the Court's reason for granting certiorari in a particular case. Two of the reasons typically given are particularly useful in determining how strongly justices feel about a case.

First, the Court sometimes notes it is granting a case "to resolve [an] important or significant question." This is a direct, though admittedly subjective, measure of how important the questions raised in a case are. To the extent this measure accurately reflects the majority author's state of mind regarding the case, opinions deciding cases granted for this reason should be written in a clearer style. Accordingly, we include a variable, *Significant Question*, equaling one if the majority opinion author noted that the case was granted to resolve an important or significant question, and zero otherwise.

Second, the Court sometimes notes it is granting a case to resolve a "conflict" between federal circuit courts or other lower courts. Claims—including by justices testifying before

Congress on the 1925 Judges' Bill—that the Court will "as a matter of course" decide on the merits all lower court conflicts are overstated (Beim and Rader 2019; Caldeira and Lempert 2020). Still, the presence of a conflict between or among lower courts is perhaps the best predictor of whether the Court will grant cert, over a broad swath of the modern Court's history (Caldeira and Lempert 2017; Caldeira and Wright 1990). Granted, this fact is subject to two interpretations. It could mean either that that justices feel obligated to grant conflict cases, regardless of interest in the underlying legal policy—or, it could mean that justices are particularly interested in cases involving conflict. However, the literature on merits decision-making in conflict cases points to the former interpretation (Bartels 2011; Lindquist and Klein 2006). As such, the implication is that in conflict cases, justices will derive relatively less benefit from writing clearly and thus be less likely to do so. We set *Conflict* equal to one if a conflict between or among lower courts is the sole stated reason for the Court granting cert, and equal to zero otherwise.

Our final variable expected to affect the benefits of writing clearly is a majority opinion author's ideological compatibility with the decision being handed down. Justices' motivation to write clearly might vary as a function of the relationship between the ideological direction of the majority opinion and their own predilections. In particular, justices who believe that opinion clarity aids implementation may be very motivated to write clearly in cases that reach outcomes aligned with their own ideological preferences but not at all so motivated in cases where a justice believes the law dictates a result contrary to their preferences. (I.e., because they care more about policy being implemented if that policy is concordant with their own ideology.Thus, we include a majority opinion author's *Ideological Compatibility* with the decision, defined as her Martin-Quinn (2002) score if the decision is conservative, and -1 times that score if the decision is liberal; the ideological directions are as classified by Spaeth et al. (2017).

Issue Area and Justice Fixed Effects

The issue area in a case may affect an opinion's clarity either by reducing costs, or by

increasing benefits. That is, certain issue areas may involve subjects that are relatively simple to write about clearly, reducing the costs of doing so. Others may be inherently more complex, increasing costs (Budziak, Hitt and Lempert 2019, 5–6, 19). Justices may also care more about some issue areas than others (Rice 2019, 116–118; Richards 2001), affecting benefits. Thus, although theory does not give directional predictions for individual issue areas, we have reason to believe that costs and/or benefits of writing clearly are affected by a case's issue area. As such, we include fixed effects for the 13 issue areas defined in Spaeth et al. (2017).[12]

To indirectly account for justice aesthetic preferences, we include fixed effects for each justice who wrote at least 50 majority opinions in our sample. (Justices who wrote fewer than 50 opinions are combined into a single "Other Justice" category.) We recognize that this does not directly measure aesthetic preferences. However, as we have argued above, aesthetic preferences are at least a plausible source of inter-justice variation. Thus, we are interested in the magnitude of inter-justice differences in clarity. If it is substantial, relative to the differences induced by case-level costs and benefits, we would have evidence that aesthetic preferences are an important factor in justice writing style, in comparison to policy-implementation related concerns.

Clarity, Readability, and Measurement

We adopt the measurement approach used in a growing body of judicial politics scholarship to operationalize the concept of clarity: the textual readability of each opinion, as measured by average word and sentence length (e.g., Black et al. 2016*b*; Black et al. 2016*a*; Goelzhauser and Cann 2014; Hansford and Coe 2019*a*; Owens, Wedeking and Wohlfarth 2013). The specific measure we use as our outcome variable is the Flesch-Kincaid Grade Level (FKGL) of a majority opinion. FKGL is defined as:

---

[12]These are Civil Procedure, Civil Rights, First Amendment, Due Process, Privacy, Attorneys, Unions, Economic Activity, Judicial Power, Federalism, Interstate Relations, Federal Taxation, and Miscellaneous.

$$0.39 \left(\frac{\text{Words}}{\text{Sentences}} \right) + 11.8 \left(\frac{\text{Syllables}}{\text{Words}} \right) - 15.59.$$

.

The formula takes into account two attributes of a text: the average number of words per sentence, and the average number of syllables per word. Thus, it is one of many similar measures that operationalize clarity as a function of word and sentence length (for one list, see Black et al. 2016b, 50). Note that as FKGL increases, clarity decreases.

A central advantage of FKGL over related measures is that it is scaled to approximate the U.S. grade level of education required to understand the text.[13] This allows for an intuitive interpretation of differences between scores—unlike other measures for which the scales are effectively arbitrary. Also, unlike some earlier readability formulas, FKGL was developed for and validated on adult readers. In the initial validation study, subjects' estimated reading comprehension ranged from grade level 6 to 16 and the texts that subjects read had FKGL scores 6.6–16.7 (Kincaid, Fishbourne, Rogers and Chissom 1975, 11–12).[14]

We next examine FKGL's connection to the concepts of clarity and readability. The concept of clarity is not self-defining, and, following Owens and Wedeking (2011), can be thought of as a multifaceted concept. Owens and Wedeking (2011) define three ways in which a judicial opinion can be clear.

First, an opinion can be *cognitively clear*, insofar as the underlying ideas in the opinion are uncomplicated. More precisely, a text reflecting a cognitively complex thought process recognizes "multiple perspectives or dimensions associated with an issue" and "relationships and connections among these perspectives or dimensions" (Owens and Wedeking 2011, 1038–1039; see also Gruenfeld 1995 and Tetlock, Bernzweig, and Gallant 1985). Owens and Wedeking (2011) proposes an automated measure of cognitive complexity based on whether the words

---

[13]The Coleman Liau Index is also similarly scaled.

[14]Subjects' reading comprehension grade level estimates were based on an independent, established test of reading comprehension.

a text fall into one of several categories defined in the text analysis program LIWC (Tausczik and Pennebaker 2010).[15]

A second type of clarity is *doctrinal clarity.* Conceptually, doctrinal clarity reflects whether a court's approach to a specific area of law over time is "stable" rather than "inconsistent" (Owens and Wedeking 2011, 1038). Of course, this kind of clarity is often considered in legal scholarship, but requires extensive (and subjective) doctrinal analysis (Owens and Wedeking 2011, 1038), making it unsuitable for large-N studies. Moreover, it is doubtful that the doctrinal clarity is a property of an individual opinion, rather than a line of opinions.

Finally, we turn to what Owens and Wedeking (2011, 1038) term *rhetorical clarity*: "how clearly-written an opinion is." Owens and Wedeking (2011) explicitly equate this concept with a text's "readability," citing FKGL and similar measures as means of operationalizing the concept. Thus, rhetorical clarity is conceptually distinct from cognitive clarity and doctrinal clarity.

Measures, like FKGL, relying on word and sentence length have been extensively validated (e.g., Kincaid et al. 1975). Most directly relevant are two studies from political scientists that validate the tie between clarity and sentence/word length in legal opinions. Hansford and Coe (2019*a*) considers how differences in linguistic complexity—equated explicitly with a lack of (rhetorical) clarity (see p. 395)—affects acceptance of Court opinions. Specifically, the authors manipulate the language of excerpts from Supreme Court opinions to produce versions that have high clarity (i.e., low FKGL score) and low clarity (i.e., high FKGL score) (Hansford and Coe 2019*b*, 36). Hansford and Coe (2019*a*, 403) finds that subjects rated the opinion variants with lower FKGL scores as both less complicated and easier to read, and read such opinions more quickly. Black et al. (2016*a*, 713–714) compares how respondents subjectively rate the clarity of Court opinion excerpts with shorter words/sentences (i.e., lower FKGL) and those with longer words/sentences. As expected, respondents rate the excerpts with shorter

[15]It is perhaps worth noting that the original coding manual for cognitive complexity explicitly warns against such an approach (Baker-Brown, Ballard, Bluck, de Vries, Suedfeld and Tetlock 1992, 22–23).

words/sentences as more clear; the authors also report that respondents retain relatively more information from such opinions.

These studies demonstrate that readers tend to describe opinions written in shorter sentences with shorter words as relatively more clear. Given the nature of the FKGL, we believe these results are best understood as confirming that readers find such opinions to have greater *rhetorical* clarity or, equivalently, readability. FKGL does not directly tap into the cognitive clarity of the ideas underlying an opinion, and—as an opinion-level measure—FKGL resists interpretation as a measure of doctrinal clarity, which is an attribute of an opinion set. It is perhaps unfortunate that the term *clarity* has been used to describe all three of these concepts, because it it contributes to a lack of definitional precision and may mean that FKGL does not line up with every reader's intuitive understanding of what clarity entails. As such, describing the measure as tapping rhetorical clarity or readability may help reduce confusion, though at the cost of terminological inconsistency with the bulk of the most directly relevant literature.[16]

This suggests that scholars of judicial politics should hesitate before deploying FKGL to test hypotheses related to the cognitive or doctrinal clarity of legal opinions. However, the measure is well-suited for testing hypotheses about the causes and consequences of opinions' rhetorical clarity (readability). The literature we are speaking to in this analysis does exactly that: it focuses how justices manipulate opinion sentence and word length to foster implementation of their favored policies by non-Court actors. Had this literature coalesced around the term *readability* or *rhetorical clarity*, we believe the only relevant change would have been

---

[16] An informal review suggests that *readability* is the more common term across disciplines, but clarity is often used too, for example in the economic literature on the text of central bank communications (e.g., Jansen 2011). Interestingly, Justice Thomas suggests yet another term, *accessibility*, to describe the idea of having a "ten-dollar idea in a five-cent sentence"—a sentence that a "parent who is not a lawyer" or a "person at the gas station" could understand (Garner 2011, 100). Thomas elaborates: "I have a wonderful buddy who is a parapalegic. Do you realize that a curb that high [showing a two- or three-inch space] is like the Great Wall of China to him in that wheelchair? Well, maybe a sentence that long is the Great Wall of China to the people who want to read about their Constitution (Garner 2011, 129)."

a somewhat increased level of terminological precision. Neither the results in that literature, nor the implications of those results, nor this paper's relationship to that literature, would be affected in any meaningful way.

Like almost every measure in the social sciences, FKGL has some noise; we do not contend that it perfectly captures the concept of clarity. Could the existence of such error explain the observed distribution of FKGL even without aesthetic preferences? We do not think so, for the following reasons.

Given normally distributed error, it is true that *some* normal distribution would obtain even if justices always wrote as clearly as possible. But *this* normal distribution is unlikely. A variance so large, given justices who are attempting to write clearly, would only obtain if the measure bore little relationship to the underlying concept of clarity—but such a claim is belied by the validation studies we have discussed. Nor is the claim consistent with the results we show below indicating substantial inter-justice variation in FKGL. Moreover, the mean of the distribution is higher than what is reasonably consistent with justices who are attempting to write as clearly as possible. One way to show this is by example. Consider Scalia's majority opinion in *California Democratic Party* v. *Jones* (530 U.S. 567) which has a FKGL that equals the sample mean to two decimal places (14.80). It contains sentences like:

> In the 1860 presidential election, if opponents of the fledgling Republican Party had been able to cause its nomination of a pro-slavery candidate in place of Abraham Lincoln, the coalition of intraparty factions forming behind him likely would have disintegrated, endangering the party's survival and thwarting its effort to fill the vacuum left by the dissolution of the Whigs.

and

> In concluding that the burden Proposition 198 imposes on petitioners' rights of association is not severe, the Ninth Circuit cited testimony that the prospect of malicious crossover voting, or raiding, is slight, and that even though the numbers

of "benevolent" crossover voters were significant, they would be determinative in only a small number of races.

These sentences are not impenetrable. Still, it is surely possible to state the ideas therein using one or more shorter, more readable sentences.

More systematically, we can compare Supreme Court opinions to a set of topically similar texts whose authors have an incentive to write clearly. Specifically, we compare our set of majority opinions and a corpus of 2000 *New York Times* stories that include the phrase "Supreme Court," published in 1980 and 1981.[17] We draw 100 samples of 100 opinions, and 100 news story segments of equal length. For every one of the 100 draws, the news story sample had the smaller FGKL mean (median difference: 1.02) and FKGL variance (median difference: 1.22). Thus, the comparable news stories are consistently written more clearly than the Court opinions.

Estimation Sample

Our sample includes 4,518 signed Supreme Court majority opinions written in October Terms 1955–2008.[18] The unit of analysis is the majority opinion. We exclude all opinions in which the Court is constrained by Congress (e.g., Owens, Wedeking and Wohlfarth 2013). This is because Owens, Wedeking and Wohlfarth (2013) argues that, in order to avoid review of their opinions by an ideologically hostile Congress, implementation-motivated justices should write opinions that are *less* clear when constrained by Congress. The authors present evidence that majority opinions become less clear as Congress becomes increasingly distant ideologically from the Court. This represents the one exception to the conventional account we can identify where writing less clearly is proposed to *enhance* the ability of justices to implement their

---

[17]1980 was the earliest year we could access *Times* stories in bulk, in NexisUni; the maximum number of downloads allowed per year (search) was 1000, so we downloaded the 1000 "most relevant" results, as determined by NexisUni, in each year.

[18]Our sample ends with the 2008 term because of data availability limitations for salience covarites.

preferred policy. Thus, to appropriately test our hypotheses, we focus on the large subset of cases where the Court is not constrained, i.e., where justices are expected to write clearly if they use writing style to aid policy implementation.[19] Specifically, we exclude from our sample cases not involving judicial review where the Court median is ideologically more extreme in a given direction than the most extreme of the House median, Senate median, and the President. To locate the relevant actors in ideological space, we use Bailey's (2007) XTI scores, which better capture the relevant ideological dimension for inter-branch relations during the Civil Rights Era than do Judicial Common Space scores (Epstein, Martin, Segal and Westerland 2007).[20]

Hypotheses and Tests

We have discussed a number of covariates that are expected to affect either the costs or the benefits of writing clear majority opinions. To briefly summarize, costs are expected to be increasing, or benefits decreasing, in the following variables—that is, these variables are expected to have a positive relationship with the dependent variable FKGL (which, recall, is itself decreasing in clarity): Multiple Issues, Workload, Total Docket, Merits Docket, Majority Coalition Diversity, and Conflict. Costs are expected to be decreasing, or benefits increasing, in the following variables—that is, these variables are expected to have a negative relationship with the outcome FKGL: Days to End of Term, Media-Based Salience, Group Salience, Legal Salience, Number of Majority Votes, Significant Question, and Ideological Compatibility.

One way to proceed would be to test hypotheses associated with each covariate, assessing whether each of the associated coefficients are substantively and statistically significant and in the expected direction. However, we are not primarily interested in the relationship between FKGL and these individual covariates (or any single covariate). Rather, we are interested in

---

[19]In our sample, the Court is constrained in about 20% of cases.

[20]For the year 2009, i.e., the latter part of the 2008 term, we use Judicial Common Space Scores, because the XTI scores end after the calendar year 2008.

whether justices write more clearly, as costs decrease and benefits increase *as a whole*. A more relevant hypothesis is then:

Hypothesis 1: In a low-cost, high-benefit scenario, justices will write majority opinions more clearly than in a high-cost, low-benefit scenario.

We make these scenarios concrete momentarily. There are several advantages to testing this hypothesis, rather then a set of hypotheses, one about each covariate. First, it speaks more directly to our question of substantive interest. Second, it allows for more straightforward and fair theory testing: supposing that we found 10 of 13 covariates are significant and in the direction predicted by the conventional theory of opinion clarity, what could we conclude? Would it be fair to claim that the theory has been falsified because a minority of the covariates did not perform as expected? The answer is at least ambiguous. It is more modest and generous to the conventional theory to assess how clarity changes with costs and benefits as a whole; modest because it does not assume that every one of our proposed covariates are actually taken into account by justices when considering costs and benefits, and generous because it allows for the conventional theory to be vindicated even if justices respond to only some of the costs and benefits. Third, it gives a single quantity—the difference in clarity between the two scenarios—that can be straightforwardly compared to estimated inter-justice differences in clarity, thereby allowing us to assess the relative importance of implementation-related and aesthetic preferences over writing style.

The low-cost/high-benefit scenario is defined as follows. Set each covariate that has a positive theoretical relationship with FKGL to its in-sample 10$^{\text{th}}$ percentile value. Set each covariate that has a negative theoretical relationship with FKGL to its in-sample 90$^{\text{th}}$ percentile value. In parallel, the high-cost/low-benefit scenario is defined by setting each covariate that has a positive theoretical relationship with FKGL to its in-sample 90$^{\text{th}}$ percentile value, and setting each covariate that has a negative theoretical relationship with FKGL to its in-sample 10$^{\text{th}}$ percentile value. The specific values are given in Table 1. For each scenario, we leave the justice and area fixed effects as observed (i.e., we average over issue areas and

justices).

Covariate	Value at Low-Cost/ High-Benefit Scenario	Value at High-Cost/ Low-Benefit Scenario
Multiple Issues	0	1
Workload	1	8
Total Docket	2296	8882
Days to End of Term	45	215
Merits Docket	87	180
Majority Coalition Diversity	2.16	3.25
Number of Majority Votes	9	5
Media-Based Salience	0.99	-0.58
Group Salience	1.18	-0.74
Legal Salience	1	0
Significant Question	1	0
Conflict	0	1
Ideological Compatibility	3.61	-1.72

Table 1. Covariate values for predicting FKGL at the low-cost/high-benefit and high-cost/low-benefit scenarios. For the former, values correspond to the 10th percentile for variables increasing in costs, and 90th percentile for variables decreasing in costs; and the 90th percentile for variables increasing in benefits, and 10th percentile for variables decreasing in benefits. For the latter, values correspond to the 90th percentile for variables increasing in costs, and 10th percentile for variables decreasing in costs; and the 10th percentile for variables increasing in benefits, and 90th percentile for variables decreasing in benefits. Justice and Issue Area fixed effects held as observed.

To test our key hypothesis, we first estimate an OLS regression predicting a majority opinion's FKGL with the set of covariates discussed above. Since the coefficients associated with individual variables are not our primary quantities of interest, we relegate the complete regression results to Table A1 in the Appendix. Briefly, we note that only three of 13 independent variables are statistically significant and in the expected direction (Conflict, Multiple Issues, and Merits Docket). Three independent variables are significant in the opposite of the direction predicted by the conventional theory. As a robustness check, we also estimate a regression that excludes the variables Majority Size and Majority Coalition Diversity as potentially endogenous (i.e., because opinion clarity could theoretically affect the size and makeup of the majority coalition). There are no relevant differences between the two regressions' estimates.

Turning to our quantity of interest, we use the regression estimates to predict the value of FKGL at a low-cost, high-benefit scenario and a high-cost, low-benefit scenario (as defined above). If justices are attempting to write clearly as a means of furthering the implementation

of their policy preferences, but are constrained by costs and/or lack of benefits, we expect the relationship posited in Hypothesis 1 to hold: Justices should write significantly more clearly (i.e., with lower FKGL) in the low-cost/high-benefit scenario than in the high-cost/low-benefit scenario. Table 2 gives these estimates and the p-value for the significance of the difference between the two scenarios.

Specification	FKGL at Low-Cost/ High-Benefit Scenario	FKGL at High-Cost/ Low-Benefit Scenario	Difference: p
Baseline (Table A1, Col 1)	14.91	15.05	.69
Limited Predictors (Table A1, Col 2)	14.92	15.03	.70

Table 2. Predicted values of FKGL in a low-cost/high-benefit scenario and in a high-cost/low-benefit scenario. See text and Table 1 for definition of each scenario. The last column gives the p-values for the significance of the differences between predicted values for each scenario. The Baseline model includes all covariates. The Limited Predictors model excludes Majority Votes and Majority Coalition Diversity as potentially endogenous. Justice and Issue Area fixed effects are set at observed values.

We find only a negligible difference between the two scenarios. In our full model, the difference between the two scenarios is less than one-seventh of a grade level. This difference is not statistically significant ($p = .69$). In the model where Majority Coalition Diversity and Majority Votes are omitted as potentially endogenous, the difference between the two scenarios is yet smaller—barely a tenth of a letter grade—and the p-value for the significance of the difference is essentially the same ($p = .7$).

Taken as a whole, these results give no support to the conventional theory of opinion clarity. There is no evidence that justices are attempting to write particularly clearly to aid policy implementation, but are constrained by costs and lack of benefits from always doing so.

How does the theory of aesthetic preferences fair? If inter-justice differences in opinion clarity are substantially larger than the difference between the two cost/benefit scenarios, that would be consistent with the idea that justices' aesthetic preferences are an important driver of opinion clarity.

Table 3 displays the pairwise FKGL comparisons between all majority opinion writers with at least 50 majority opinions, based on 22 iterations of Model 1, Table A1, each with a different justice as the excluded category. This table depicts the statistically significant differences in opinion clarity among justice-pairs (holding constant all other covariates). The justice whose initials appear at the beginning of each row serves as the excluded category for justice fixed effects, for a given model iteration. The justice whose initials appear at the top of each column is the justice being compared to the excluded justice in a given iteration. Any cell including a coefficient represents a statistically significant difference in clarity between the row justice and the column justice. Negative coefficients indicate that the justice identified in the column writes more clearly than the justice identified in the row. Positive coefficients indicate that the column justice writes less clearly than the row justice. For example, ceteris paribus, Justice Douglas writes 1.7 grade levels lower (more clearly) than Justice Rehnquist. The cell entry '.' indicates nonsignificant differences in FKGL between justices. For example, the difference in clarity between Justices Rehnquist and Brennan is not statistically significant.

	ak	as	bw	ct	ds	ew	ff	hab	hlb	jh	js	lp	ps	rg	sb	so	tc	tm	web	wjb	wd	wr
ak		1.1	0.5	.	1.9	.	0.5	-0.9	.	1.1	.	-1.0	-0.4	-0.5	.	.	-0.6	.	.	.	-1.3	0.4
as	-1.1		-0.7	-0.9	0.8	-1.2	-0.6	-2.0	-1.2	.	-1.0	-2.2	-1.5	-1.7	-1.3	-1.2	-1.8	-1.2	-1.2	-0.8	-2.5	-0.8
bw	-0.5	0.7		.	1.5	-0.5	.	-1.3	-0.5	0.7	-0.3	-1.5	-0.8	-1.0	-0.6	-0.5	-1.1	-0.6	-0.6	.	-1.8	.
ct	.	0.9	.		1.7	.	.	-1.1	.	0.9	.	-1.3	-0.6	-0.8	.	.	-0.9	.	.	.	-1.6	.
ds	-1.9	-0.8	-1.5	-1.7		-2.0	-1.4	-2.8	-2.0	-0.8	-1.8	-3.0	-2.3	-2.5	-2.1	-2.0	-2.6	-2.0	-2.0	-1.6	-3.3	-1.6
ew	.	1.2	0.5	.	2.0		0.6	-0.8	.	1.2	.	-0.9	.	.	.	.	-0.6	.	.	0.4	-1.2	0.4
ff	-0.5	0.6	.	.	1.4	-0.6		-1.4	-0.6	0.6	.	-1.6	-0.9	-1.1	-0.7	-0.6	-1.2	-0.6	-0.6	.	-1.8	.
hab	0.9	2.0	1.3	1.1	2.8	0.8	1.4		0.8	2.0	1.1	.	0.5	.	0.7	0.8	.	0.8	0.8	1.2	-0.4	1.3
hlb	.	1.2	0.5	.	2.0	.	0.6	-0.8		1.2	.	-1.0	.	.	.	.	-0.6	.	.	0.4	-1.2	0.4
jh	-1.1	.	-0.7	-0.9	0.8	-1.2	-0.6	-2.0	-1.2		-1.0	-2.2	-1.5	-1.7	-1.3	-1.2	-1.8	-1.2	-1.2	-0.8	-2.5	-0.8
js	.	1.0	0.3	.	1.8	.	.	-1.1	.	1.0		-1.2	-0.5	-0.7	.	.	-0.8	-0.3	-0.3	.	-1.5	.
lp	1.0	2.2	1.5	1.3	3.0	0.9	1.6	.	1.0	2.2	1.2		0.6	0.5	0.8	1.0	0.4	0.9	0.9	1.3	.	1.4
ps	0.4	1.5	0.8	0.6	2.3	.	0.9	-0.5	.	1.5	0.5	-0.6		.	.	0.3	.	0.3	0.3	0.7	-0.9	0.7
rg	0.5	1.7	1.0	0.8	2.5	.	1.1	.	.	1.7	0.7	-0.5	.		.	0.5	.	0.4	0.4	0.8	-0.8	0.9
sb	.	1.3	0.6	.	2.1	.	0.7	-0.7	.	1.3	.	-0.8	.	.		.	.	.	.	0.5	-1.1	0.6
so	.	1.2	0.5	.	2.0	.	0.6	-0.8	.	1.2	.	-1.0	-0.3	-0.5	.		-0.6	.	.	0.4	-1.3	0.4
tc	0.6	1.8	1.1	0.9	2.6	0.6	1.2	.	0.6	1.8	0.8	-0.4	.	.	.	0.6		0.5	0.5	0.9	-0.7	1.0
tm	.	1.2	0.6	.	2.0	.	0.6	-0.8	.	1.2	0.3	-0.9	-0.3	-0.4	.	.	-0.5		.	0.4	-1.2	0.5
web	.	1.2	0.6	.	2.0	.	0.6	-0.8	.	1.2	0.3	-0.9	-0.3	-0.4	.	.	-0.5	.		0.4	-1.2	0.5
wjb	.	0.8	.	.	1.6	-0.4	.	-1.2	-0.4	0.8	.	-1.3	-0.7	-0.8	-0.5	-0.4	-0.9	-0.4	-0.4		-1.6	.
wd	1.3	2.5	1.8	1.6	3.3	1.2	1.8	0.4	1.2	2.5	1.5	.	0.9	0.8	1.1	1.3	0.7	1.2	1.2	1.6		1.7
wr	-0.4	0.8	.	.	1.6	-0.4	.	-1.3	-0.4	0.8	.	-1.4	-0.7	-0.9	-0.6	-0.4	-1.0	-0.5	-0.5	.	-1.7	

Table 3. Each row represents an iteration of Model 1, Table A1. The justice whose initials lead a given row is the justice who serves as the excluded category for justice-level fixed effects in that iteration. Justice initials heading each column represent the justice being compared to the excluded category. Cells with numerical values indicate a statistically significant difference in opinion clarity between the column and row justices. The coefficient indicates the FKGL grade level difference (controlling for all other covariates in Model 1, Table A1). Positive coefficients indicate that the row justice writes more clearly than the column justice. Negative coefficients indicate that the row justice writes less clearly than the column justice. For example, Douglas (wd) is the clearest writer in our sample. Cells with a '.' indicate nonsignificant differences between the two justices. ak=Kennedy; as=Scalia; bw=White; ct=Thomas; ds=Souter; ew=Warren; hab=Blackmun; hlb=Black; jh=Harlan; js=Stevens; lp=Powell; ps=Stewart; rg=Ginsburg; sb=Breyer; so=O'Connor; tc=Clark; tm=Marshall; web=Burger; wjb=Brennan; wd=Douglas; wr=Rehnquist.

The pattern that emerges is clear: holding costs, benefits, and issue areas constant, justices differ significantly from each other with respect to opinion clarity. Of the 231 pairwise comparisons, 168 (73%) demonstrate statistically significant differences in FKGL. Not only is the number of significant differences noteworthy, so are the magnitudes of the differences. Many differences are greater than a full grade level, several are greater than two grade levels, and a few as large as three grade levels. For *every* justice, the mean magnitude of the inter-justice differences is at least three times larger than the magnitude of the difference between the low-cost/high-benefit and high-cost/low-benefit scenarios under the Baseline Model (.14, see Table 2). At the extreme, William Douglas writes nearly three and a half grade levels more clearly than David Souter. This of course accords with qualitative accounts of those justices' writing style, including those we described above (Rosen 1993; Rosen 2009; Small 2007). Generally speaking, we show substantial inter-justice variation, and can plausibly attribute that variation to justices' aesthetic preferences.

Discussion

We started our analysis with the observation that there is a wide range of clarity in Court majority opinions, which is unexpected in a world where justices are mostly concerned about policy implementation and use opinion clarity to achieve that goal. We considered whether justices may be *attempting* to write clearly, but are constrained from doing so by case-level factors. However, we did not find evidence that case-level costs or benefits systematically and substantially affect the level of clarity in Court majority opinions. Thus, we cannot say that, in general, justices are targeting clarity, but are constrained from achieving it.

We also proposed that justices' aesthetic preferences over clarity, unrelated to policy implementation concerns, could theoretically account for the observed distribution of opinion clarity. We have argued that aesthetic preferences are in fact important to justices, and that they are likely to vary over justices. Empirically, we found that there is substantial inter-justice variation in opinion clarity, even accounting for case-level costs and benefits, and case

issue areas. As such, we conclude that aesthetic preferences are a likely, or at least plausible, explanation for why justices write as they do.

It is true the empirical evidence for our theory of aesthetic preferences is indirect. Ideally, we would have a direct measure of w_j^*, a justice's aesthetically most-preferred level of clarity. But the prospects for obtaining such a measure are somewhat dim. Using a justice's majority opinions to derive such a measure appears fraught with endogeneinty whose specific forms would be difficult to diagnose, let alone cure. And using separate (concurring and dissenting) opinions or off-bench writings seems to require—at a minimum—the assumption that justices' aesthetic preferences are constant across majority opinions and other texts. Thus, indirect tests may be the best we can do, though we certainly encourage scholars to derive feasible direct measures.

Are there other plausible explanations for inter-justice variation in clarity besides aesthetic preferences? One possibility we can foreclose is that any sizeable subset of justices are consistently more policy-motivated than others. A justice-by-justice analysis analogous to the one in Table 2 shows that only two justices, Blackmun and Douglas, write significantly ($p < .05$) more clearly in the low-cost/high-benefit scenario compared to the high-benefit/low-cost scenario. (This analysis is included with the replication code.) Thus, we have, at most, two out of 22 justices who show appreciable signs of modifying their writing style for reasons of policy implementation.

Nor are there obvious temporal trends that account for the differences. All else equal, the five clearest writers are Douglas, Powell, Blackmun, Clark, and Ginsburg, and the five least clear, Souter, Scalia, Harlan, White, and Rehnquist.[21] In sum, then, we think aesthetic preferences are the most likely explanation for the inter-justice variation we see, but we acknowledge that aesthetic preferences probably do not account for all of the inter-justice variation.

Our research thus fits with a line of studies that suggest that non-policy-related preferences significantly shape behavior on the Court (e.g., Baum and Devins 2019; Maltzman and

---

[21]In a specification included with the replication code, we add fixed effects for decade. This yields no obvious time trends and does not otherwise affect our results.

Wahlbeck 2004; Owens and Wohlfarth 2019). On the narrower question of opinion writing style, our results indicate that while justices may employ opinion clarity to achieve policy implementation related goals in certain circumstances (e.g., Black et al. 2016*b*), the strategic use of clarity is not a general phenomenon. Also, our results suggest more weight for aesthetic preferences, relative to policy-motivated goals, than the conventional theory of opinion clarity would allow. This implies that studies of writing style should account for aesthetic preferences, at least by including justice fixed effects. We suggest several avenues open to research: directly measuring aesthetic preferences (over clarity or otherwise), explaining the source of these preferences, and examining their consequences—whether more directly in terms of policy implementation or in terms of audience reception more broadly.

References

Bailey, Michael A. 2007. "Comparable Preference Estimates Across Time and Institutions for the Court, Congress, and Presidency." *American Journal of Political Science* 51(3): 433–448.

Baird, Vanessa A. 2004. "The Effect of Politically Salient Decisions on the US Supreme Court's Agenda." *Journal of Politics* 66(3): 755–772.

Baker-Brown, Gloria, Elizabeth J. Ballard, Susan Bluck, Brian de Vries, Peter Suedfeld and Philip E. Tetlock. 1992. "Coding Manual for Conceptual/Integrative Complexity." Typescript. On File with Author.

Bartels, Brandon L. 2011. "Choices in Context: How Case-Level Factors Influence the Magnitude of Ideological Voting on the U.S. Supreme Court." *American Politics Research* 39(1): 142–175.

Baum, Lawrence. 1997. *The Puzzle of Judicial Behavior.* Ann Arbor, MI: University of Michigan Press.

Baum, Lawrence. 2006. *Judges and Their Audiences: A Perspective on Judicial Behavior.* Princeton, NJ: Princeton University Press.

Baum, Lawrence and Neal Devins. 2019. *The Company They Keep: How Partisan Divisions Came to the Supreme Court.* New York: Oxford University Press.

Beim, Deborah and Kelly Rader. 2019. "Legal Uniformity in American Courts." *Journal of Empirical Legal Studies* 16(3): 448–478.

Black, Ryan C., Ryan J. Owens, Justin Wedeking and Patrick C. Wohlfarth. 2016a. "The Influence of Public Sentiment on Supreme Court Opinion Clarity." *Law and Society Review* 50(3): 703–732.

Black, Ryan C., Ryan J. Owens, Justin Wedeking and Patrick C. Wohlfarth. 2016*b*. *Supreme Court Opinions and Their Audiences.* Cambridge, UK: Cambridge University Press.

Box-Steffensmeier, Janet and Dino P. Christenson. 2012. "Database on Supreme Court Amicus Curiae Briefs." Version 1.0 [Computer File]. Available at: `amicinetworks.com`.

Budziak, Jeffrey, Matthew P. Hitt and Daniel Lempert. 2019. "Determinants of Writing Style on the United States Circuit Courts of Appeal." *Journal of Law and Courts* 7(1): 1–28.

Caldeira, Gregory A. and Daniel Lempert. 2017. "Agenda Control in the Hughes Court, OT 1939." Presented at 2017 Annual Meeting of the Midwest Political Science Association.

Caldeira, Gregory A. and Daniel Lempert. 2020. "Selection of Cases for Discussion: The U.S. Supreme Court, OT 1939, 1968, and 1982." *Journal of Law and Courts* 8(2): 381–395.

Caldeira, Gregory A. and John R. Wright. 1990. "The Discuss List: Agenda Building in the Supreme Court." *Law and Society Review* 24(3): 807–836.

Caldeira, Gregory and John Wright. 1988. "Organized Interests and Agenda Setting at the U.S. Supreme Court." *American Political Science Review* 82(4): 1109–1127.

Cardozo, Benjamin N. 1931. *Law and Literature and Other Essays and Addresses.* New York: Harcourt Brace and Co.

Carlson, Keith, Michael A. Livermore and Daniel Rockmore. 2016. "A Quantitative Analysis of Writing Style on the U.S. Supreme Court." *Washington University Law Review* 93(6): 1461–1510.

Clark, Tom S., Jeffrey R. Lax and Douglas Rice. 2015. "Measuring the Political Salience of Supreme Court Cases." *Journal of Law and Courts* 3(1): 37–65.

Collins, Paul M. 2008. *Friends of the Supreme Court: Interest Groups and Judicial Decision Making.* Oxford, UK: Oxford University Press.

Dalvean, Michael and Galbadrakh Enkhbayar. 2018. "Assessing the Readability of Fiction: A Corpus Analysis and Readability Ranking of 200 English Fiction Texts." *Linguistic Research* 35(Special Edition): 137–170.

Dodson, Scott and Ami Dodson. 2015. "A Top Ten Ranking of the Supreme Court Literary Justice." *Green Bag* 18(2): 429–434.

Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics, and Organizations* 23(2): 303–325.

Epstein, Lee and Jack Knight. 1998. *The Choices Justices Make.* Oxford, UK: Oxford University Press.

Federal Judicial Center. 2013. *Judicial Writing Manual: A Pocket Guide for Judges.* 2nd ed. Washington, DC: Federal Judicial Center.

Federal Judicial Center. 2020. *Law Clerk Handbook: A Handbook for Law Clerks to Federal Judges.* 4th ed. Washington, DC: Federal Judicial Center.

Frankenreiter, Jens. 2019. Writing Style and Legal Traditions. In *Law as Data: Computation, Text, and the Future of Legal Analysis*, ed. Michael A. Livermore and Daniel N. Carlson. Santa Fe Institute Press pp. 153–192.

Garner, Bryan A. 2011. "Interviews With Supreme Court Justices." *The Scribes Journal of Legal Writing* 13(1): 5–182.

Goelzhauser, Greg and Damon M. Cann. 2014. "Judicial Independence and Opinion Clarity on State Supreme Courts." *State Politics and Policy Quarterly* 14(2): 123–141.

Gruenfeld, Deborah H. 1995. "Status, Ideology, and Integrative Complexity on the U.S. Supreme Court: Rethinking the Politics of Political Decision Making." *Journal of Personality and Social Psychology* 68(1): 5–20.

Hansford, Thomas G. and Chelsea Coe. 2019*a*. "Linguistic Complexity, Information Processing, and Public Acceptance of Supreme Court Decisions." *Political Psychology* 40(2): 395–412.

Hansford, Thomas G. and Chelsea Coe. 2019*b*. "Supporing Information for "Linguistic Complexity, Information Processing, and Public Acceptance of Supreme Court Decisions"."". Downloaded from: https://onlinelibrary.wiley.com/doi/abs/10.1111/pops.12497.

Jansen, David-Jan. 2011. "Does the Clarity of Central Bank Communications Affect Volatility in Financial Markets? Evidence From Humprhey-Hawkins Testimonies." *Contemporary Economic Policy* 29(4): 494–509.

Kincaid, J. Peter, Robert P. Fishbourne, Jr., Richard L. Rogers and Brad S. Chissom. 1975. "Derivation of New Readability Formulas for Navy Enlisted Personnel." Naval Technical Training Branch Research Branch Report 8-75.

Krewson, Christopher N. 2019. "Strategic Sensationalism: Why Justices Use Emotional Appeals in Supreme Court Opinions." *Justice System Journal* 40(4): 319–336.

Lindquist, Stefanie A. and David E. Klein. 2006. "The Influence of Jurisprudential Considerations on Supreme Court Decisionmaking: A Study of Conflict Cases." *Law & Society Review* 40(1): 135–162.

Maltzman, Forrest, James F. Spriggs and Paul J. Wahlbeck. 2000. *Crafting Law on the Supreme Court: The Collegial Game.* Cambridge, UK: Cambridge University Press.

Maltzman, Forrest and Paul J. Wahlbeck. 2004. "A Conditional Model of Opinion Assignment on the U.S. Supreme Court." *Political Research Quarterly* 57(4): 551–563.

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10(2): 134–153.

Niemi, Richard G. and Larry M. Bartels. 1985. "New Measures of Issue Salience: An Evaluation." *Journal of Politics* 47(4): 1212–1220.

Owens, Ryan J. and Patrick C. Wohlfarth. 2019. "The Influence of Home-State Reputation and Public Opinion on Federal Circuit Court Judges." *Journal of Law and Courts* 7(2): 187–214.

Owens, Ryan and Justin Wedeking. 2011. "Justices and Legal Clarity: Analyzing the Complexity of U.S. Supreme Court Opinions." *Law and Society Review* 45(4): 1027–1061.

Owens, Ryan, Justin Wedeking and Patrick Wohlfarth. 2013. "How the Supreme Court Alters Opinion Language to Evade Congressional Review." *Journal of Law and Courts* 1(1): 35–59.

Posner, Richard. 1993. "What do Judges and Justices Maximize?" *Supreme Court Economic Review* 3(1): 1–41.

Posner, Richard A. 1995. "Judges' Writing Styles (And Do They Matter?)." *University of Chicago Law Review* 62(4): 1421–1449.

Rice, Douglas. 2019. "Measuring the Issue Content of Supreme Court Opinions." *Journal of Law and Courts* 7(1): 107–127.

Richards, Neil M. 2001. "The Supreme Court Justice and "Boring" Cases." *The Green Bag* 4(4): 401–407. 2nd Series.

Rosen, Jeffrey. 1993. "Poetic Justice." *The New Republic* 208(10): 25–27.

Rosen, Jeffrey. 2009. "The Stealth Justice." *New York Times* May 2.

Rosenthal, Jeffrey S. and Albert H. Yoon. 2011. "Detecting Multiple Authorship of U.S. Supreme Court Legal Decisions Using Function Words." *Annals of Applied Statistics* 5(1): 283–308.

Small, Marshall L. 2007. "William O. Douglas Remembered: A Collective Memory by WOD's Law Clerks." *Journal of Supreme Court History* 32(3): 297–334.

Spaeth, Harold J., Lee Epstein, Jeffrey A. Segal, Ted Ruger, Andrew D. Martin and Sarah Benesh. 2017. "The Supreme Court Database." Available at: `scdb.wustl.edu`.

Tausczik, Yla R. and James W. Pennebaker. 2010. "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods." *Journal of Language and Social Psychology* 29(0): 24–54.

Tetlock, Philip E., Jane Bernzweig and Jack L. Gallant. 1985. "Supreme Court Decision Making: Cognitive Style as Predictor of Ideological Consistency of Voting." *Journal of Personality and Social Psychology* 48(5): 1227–1239.

Vance, Ruth C. 2011. "Judicial Opinion Writing: An Annotated Bibliography." *Legal Writing* 17(1): 197–237.

Varsava, Nina. 2021. "Professional Irresponsibility and Judicial Opinions." Downloaded from: https://ssrn.com/abstract=3825848. Forthcoming, Houston Law Review.

Wald, Patricia M. 1995. "A Reply to Judge Posner." *The University of Chicago Law Review* 62(4): 1451–1454.

Appendix

Covariate	Coefficient (1)	Coefficient (2)
Media Based Salience $(-)$	0.011	0.011
	(0.037)	(0.037)
Group Salience $(-)$	0.091*	0.091*
	(0.023)	(0.023)
Legal Salience $(-)$	−0.051	−0.051
	(0.065)	(0.065)
Significant Question $(-)$	0.346*	0.346*
	(0.073)	(0.073)
Conflict $(+)$	0.245*	0.244*
	(0.062)	(0.062)
Ideological Compatibility $(-)$	0.001	0.001
	(0.010)	(0.010)
Multiple Issues $(+)$	0.201*	0.201*
	(0.065)	(0.065)
Workload $(+)$	−0.003	−0.003
	(0.010)	(0.010)
Total Docket (/100) $(+)$	−0.006*	−0.006*
	(0.002)	(0.002)
Time to End of Term (/100) $(-)$	−0.037	−0.037
	(0.038)	(0.038)
Merits Docket (/100) $(+)$	0.581*	0.582*
	(0.132)	(0.131)
Majority Coalition Diversity $(+)$	0.001	
	(0.042)	
Number of Majority Votes $(-)$	−0.003	
	(0.016)	
Justice and Issue Area FEs	X	X
N	4518	4518

Table A1. Dependent variable: FKGL. Expected coefficient sign in parentheses after covariate name. OLS coefficients; standard errors below in parentheses. Fixed effects and constant not shown. Total Docket, Time to End of Term, and Merits Docket are divided by 100 (equivalently, the coefficients and standard errors are multiplied by 100) for presentational purposes. (*: $p < 0.05$.)