AN OUTCOMES ANALYSIS OF SCOPE OF REVIEW STANDARDS

PAUL R. VERKUIL*

TABLE OF CONTENTS

* PROFESSOR OF LAW, CARDOZO LAW SCHOOL, YESHIVA UNIVERSITY; VISITING PROFESSOR, UNIVERSITY OF PENNSYLVANIA LAW SCHOOL, 2001-2002; PRESIDENT EMERITUS, THE COLLEGE OF WILLIAM AND MARY. I WOULD LIKE TO THANK MEMBERS OF THE PENN LAW FACULTY FOR MANY HELPFUL COMMENTS AT A WORKSHOP IN APRIL 2002 AND ALSO JOHN DUFFY, MICHAEL HERZ, JEFF LUBBERS, DAN MELCALFE, ED RUBIN, AND DAVID VLADECK FOR THEIR COMMENTS ON EARLIER DRAFTS. I AM GRATEFUL FOR THE ASSISTANCE OF MATT FORMAN, CARDOZO LAW 2001, AND OWEN KENDLER, PENN LAW 2002.

Under existing standards, then, the courts may narrow their review to satisfy the demands for administrative discretion, and they may broaden it close to the point of substituting their judgment for that of the administrative agency.[1]

After fifty years ... we have yet to agree on how this review should operate in practice. We are still struggling with where to draw the line between obsequious deference and intrusive scrutiny.[2]

INTRODUCTION

1. REPORT OF THE COMM. ON ADMIN. PROC., APPOINTED BY THE ATT'Y GEN., AT THE REQUEST OF THE PRESIDENT, TO INVESTIGATE THE NEED FOR PROCEDURAL REFORM IN VARIOUS ADMIN. TRIBUNALS AND TO SUGGEST IMPROVEMENTS THEREIN, S. DOC. NO. 77-8, at 91 (1941) [hereinafter 1941 ADMIN. PROC. REPORT]. This report has long been hailed for its insights into the administrative process.

2. Patricia M. Wald, *Judicial Review in Midpassage: The Uneasy Partnership Between Courts and Agencies Plays On*, 32 TULSA L.J. 221, 258 (1996).

The purpose of calibrating the breadth—or scope—of judicial review over fact finding by administrative agencies is ultimately to allocate decision-making responsibility between the executive and judicial branches. Because Congress usually makes these decisions, all three branches have a stake in the process. In assigning oversight responsibilities, Congress makes a choice: it weighs the desire for efficient and timely agency action against the need to ensure consistent and fair decision making. In balancing these considerations, Congress intends factual support for agency decisions to be subject to varying levels of scrutiny or, on occasion, to be free from scrutiny.[3] Straightforward enough, one would think. Yet, as the introductory quotes suggest, after all these years, reviewing judges are still struggling to make sense of these standards, especially as they apply to scope of review of facts or of law and policy.[4]

I. SCOPE OF REVIEW

It is doubtful that Congress wants scope of review to be an irrelevant labeling exercise. Instead, one might reasonably expect that Congress wants outcomes, defined in terms of affirmances, remands, and reversals of agency actions, to vary according to the scope of review standard chosen (or at least to find some judicial recognition of these expectations). But it seems the outcomes question is rarely asked and its premise remains unexamined.

To explore the relationship of outcomes to standards, this Article makes a preliminary attempt to measure outcomes against the relevant

3. When agency decisions are free from scrutiny, they are not subject to judicial review; in terms of this Article, their affirmance rate would be 100%. Situations of agency nonreviewability are infrequent and disfavored. *See, e.g.*, Yakus v. United States, 321 U.S. 414, 433 (1944) (denying judicial review in criminal enforcement proceedings). *But see* Adamo Wrecking Co. v. United States, 434 U.S. 275, 290 (1978) (Powell, J., concurring) (criticizing *Yakus* as "ambiguous" and distinguishing it as a wartime measure); *see also* Paul R. Verkuil, *Congressional Limitations on Judicial Review of Rules*, 57 TUL. L. REV. 733 (1983). The subject of nonreviewability is beyond the scope of this Article, except to note that Congress' instructions in this regard are still followed by the courts unless the Constitution dictates otherwise. *See* INS v. St. Cyr, 533 U.S. 289, 300 (2001) ("A construction of the amendments at issue that would entirely preclude review of a pure question of law by any court would give rise to substantial constitutional questions.").

4. Scope of review of questions of law and policy are also a challenging exercise, but it is at least one that has received extensive judicial analysis. *See infra* notes 28, 86, and accompanying text (discussing the *Chevron* doctrine). By contrast, scope of review over facts, including application of law to facts, remains a neglected activity.

scope of review provisions. This "outcomes analysis" pro-duces some intriguing correlations between results and formulas, along with some surprises that themselves serve to raise more questions. Although such an exercise can never produce total agreement about how scope of review standards should operate, it facilitates a better understanding of why Congress differentiates among these standards and why the lesson sometimes is lost on the courts.

A. The Art of Scope of Review

Think of the word "scope" in "scope of review" as a contraction of "telescope." Like a telescope, scope of review offers either a narrow aperture to limit the breadth of judicial scrutiny, thereby increasing the area of agency discretion, or a wider lens to expand judicial oversight, thereby decreasing agency discretion. Once Congress supplies the lens,[5] the courts and agencies must try to bring Congress' intended level of judicial scrutiny into focus. This exercise inevitably produces margins that are fuzzy and obscure. After all, the object being observed is not easy to contemplate: instead of viewing a beautiful sunset, the judge's eyes are cast upon a mind-numbing pile of documents.

Moreover, unlike, say, a National Football League official uti-lizing "instant replay" to review a play challenged on the field, the reviewing judge is looking at an event that occurred years earlier which has few clear guideposts. When "field judges" are told by the National Football League not to reverse a play unless the call is "clearly wrong," the process might be expected to generate a fairly predictable reversal rate.[6]

---

5. Of course, Congress does not always supply a lens. Statutes are sometimes silent on scope of review, or indeed on review at all. In those situations, the courts, by adopting a presumption of reviewability, will provide the standard of review. The 1941 Attorney General's Report on Administrative Procedure explains:

> Like the area of judicial review, the extent to which administrative action within that area will be subjected to judicial scrutiny is also largely determined by court decisions. The courts developed standards as to the scope of judicial review when legislation did not provide them. To a large extent, subsequent legislation dealing with the matter has either enacted the judicially formulated standards or has been so interpreted by the courts that no difference resulted.

1941 ADMIN. PROC. REPORT, *supra* note 1, at 87. The classic judicial articulation of the presumption of reviewability is *Abbott Labs v. Gardner*, 387 U.S. 136, 140-41 (1967).

6. An estimated thirty-five percent of the reviewed calls are reversed. *See* Mike Sando, *Return of Instant Replay Not Off to Good Start*, NEWS TRIB. (Tacoma, Wash.), Nov. 28, 1999, at C4. But even National Football League field judges have varying "reversal rates" that must be calculated

But even in the limited world of sports not all calls are automatic. In baseball, for example, umpires regularly make controversial calls with far less consistency than one might imagine.[7] Few reviewing functions, it seems, are routine, automatic, or bland.[8]

---

by coaches when they decide to risk a challenge, the failure of which can cost them an often crucial time-out. *See* Jarrett Bell, *Replay Will be Strategy: Challenges Become Coach's Tool*, USA TODAY, Mar. 18, 1999, at C3.

7. Concerning the newly revised strike zone, one commentator observed: "The reality is no one even knows if umpires will call pitches the way they have been instructed." Murray Chass, *Calling Pitches By the Book Won't Be Easy*, N.Y. TIMES, Mar. 30, 2001, at D1. In an effort to speed up the game, Major League Baseball ordered umpires to call more strikes, or more precisely, to limit the total pitch count to 270 per game. *See* Murray Chass, *Call More Strikes, Umpires are Told*, N.Y. TIMES, July 14, 2001, at D1. Responding to the ensuing outcry by purists, within days Major League Baseball reversed course and agreed not to use pitch counts to measure umpire performance, and the controversy dissipated. *See* Murray Chass, *Baseball Retreats in Dispute Over Umpires' Pitch Counts*, N.Y. TIMES, July 19, 2001, at D1. The issue remains who should decide what the strike zone is, i.e., who defines "what is good for the game." *See infra* Part III.A. (comparing this issue to the Social Security Administration's (SSA) efforts to control Administrative Law Judge (ALJ) reversal rates and average caseloads).

8. Perhaps the court in *Scenic Hudson Preservation Conference v. Federal Power Commission*, 354 F.2d 608 (2d Cir. 1965), erred when it famously refused to permit administrative agencies to behave like umpires "blandly calling balls and strikes." *Id.* at 620 (refusing to let the Federal Power Commission (FPC) play the seemingly passive role of an umpire in dealing with the environment).

Unlike the field judge, the appellate judge does not get freeze frames of the action below. And, unlike the field judge and the umpire,[9] the appellate judge does not get to view the same situation over and over again. Instead, he or she is presented with a constantly changing stack of papers that help advocates reconstruct the action below. Depending on the type of agency action under review, this record may be formal and contain a written decision with transcripts, exhibits, briefs, and other submissions, or it may be informal, and contain things like letters or notations rejecting a request or imposing obligations. In most circumstances, the appellate judge lacks a clear view of the action below.

Once the court receives the agency record and hears arguments, it must apply the relevant scope of review standard. But scope is one dimension, intensity another. Although Congress usually es-tablishes scope, the Supreme Court traditionally determines how close or hard the courts must look in a given situation. The Court's directions are often complicated and sometimes inconsistent. For example, one well-known formulation combines a "narrow" scope of review standard with a "searching" inquiry.[10] And sometimes the Court infuses the arbitrary and capricious standard with a "hard look" requirement.[11]

---

9. Umpires are given a defined strike zone but often have difficulty locating it consistently, at least according to players and managers. In this circumstance it is consistency that is sought, not some manifestation of the ideal strike zone. *See* Dave Anderson, *The Poison Threatening The Umpires*, N.Y. TIMES, July 18, 2001, at D1 (quoting Ted Williams, who "once described consistency as 'the one necessary ingredient' for a home plate umpire").

10. In *Citizens to Preserve Overton Park, Inc. v. Volpe*, 401 U.S. 402 (1971), the Court rejected the application of the substantial evidence test and de novo review in favor of the less stringent arbitrary and capricious review. Even under this narrow standard, however, the Court required "probing, in depth review." *Id.* at 415. *But see* 2 KENNETH CULP DAVIS & RICHARD J. PIERCE, JR., ADMINISTRATIVE LAW TREATISE § 11.5 (3d ed. 1994) (arguing that "the *Overton Park* opinion

overstated the requirements of the arbitrary and capricious test").

11. *See, e.g.*, Motor Vehicle Mfrs. Ass'n v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 42-44 (1983).

These instructions test a reviewer's mental and visual dexterity. The judge must be an omniscient observer but not an omnivorous decision maker. In effect, the reviewer often must see, but yet suspend judgment.[12] This tension forces a judge to live in a dual reality, not unlike what jurors must do when a trial judge instructs them to disregard what they have seen or heard.[13] Presumably reviewing judges are more adept at this task than jurors, but it still requires a talent for the interpretative role. This is one reason why judicial review is more art than science.

B. The Administrative Procedure Act as a Guide

Mental balancing acts for reviewing judges come with the territory. In drafting the Administrative Procedure Act (APA) scope of review provisions Congress sought to bring order to the oversight function.[14]

---

12. In a way the reviewing judge is being asked to perform the literary task of engaging in a "willing suspension of disbelief." That phrase, coined by Samuel Coleridge, asks the reader to transfer from his "inward nature" a "semblance of truth" to sustain the poetic moment. *See* J.A. CUDDON, A DICTIONARY OF LITERARY TERMS AND LITERARY THEORY 1044 (3d ed. 1991).

13. Stanley Sue et al., *Effects of Inadmissible Evidence on the Decisions of Simulated Jurors: A Moral Dilemma*, 3 J. APPLIED SOC. PSYCHOL. 345, 351-52 (1973) (analyzing jury simulation research and finding that judgments of mock jurors are often influenced by evidence the judge instructs them to disregard).

14. *See* Administrative Procedure Act, Pub. L. No. 404, ch. 324, §§ 1-12, 60 Stat. 237 (1946)

In *Universal Camera Corp. v. NLRB*,[15] the Court struggled to infuse the APA's substantial evidence test with predictive powers. Justice Frankfurter instructed the circuit courts uniformly to consider the fact finding role of the hearing examiner,[16] but he remained skeptical of judges' ability to carry out this instruction. Judges, he memorably warned, were not "automata."[17] In effect, he discounted in advance the probability that instructions on scope of review, even instructions sent twice to the best circuit in the land, would yield consistent results.

(codified as amended in scattered sections of 5 U.S.C.).

15. 340 U.S. 474 (1951).

16. In *Universal Camera*, the Court reversed the Second Circuit's application of the substantial evidence standard for inadequately crediting the hearing examiner's conclusions on witness credibility. *Id*. at 496. On remand, Judge Frank sought to correct the analysis by distinguishing between primary inferences of fact, such as witness demeanor, and secondary inferences of fact. NLRB v. Universal Camera Corp., 190 F.2d 429, 432 (2d Cir. 1951) (Frank, J., concurring).

17. *Universal Camera*, 340 U.S. at 489. Of course, not all scholars feel that the scope of review standards are unworkable. *See* DAVIS & PIERCE, *supra* note 10, § 11.2 (noting that substantial evidence review "has been among the most stable and satisfactory features of our system of administrative law").

Scope of review remains unpredictable and contentious. Consider the recent decision in *Easley v. Cromartie*,[18] in which the Supreme Court split five-to-four regarding the meaning and application of the clearly erroneous standard.[19] Because clearly erroneous is an even more familiar standard than substantial evidence, one wonders how the Justices can divide so evenly over its application. Fifty years after *Universal Camera*, the Supreme Court still has great difficulty leading the way.

Once the Court put aside any theory of mechanical application, scope of review doctrine seems to have suffered from benign neglect. Rather than locating guideposts between mechanical application and freelancing, courts and commentators have largely abandoned the field.[20] With Congress unsure of what to do next, all options remain open. Evaluation of outcomes may help to sharpen differ-ences among the various conceptually distinct but empirically muddled review standards and, by so doing, regain doctrinal clarity. The Court has acted in related contexts where outcomes have been affected through judicially controlled guidelines, such as scope of review over questions of law and policy.[21]

---

18. 532 U.S. 234 (2001).

19. In *Cromartie*, the Court divided over the question whether the three judge panel's decision to set aside North Carolina's legislative redistricting boundaries was a violation of the "clearly erroneous" standard of review when race was a "predominant" motive in creating the districts. *Id*. at 259-67 (Thomas, J., dissenting).

20. Perhaps the best (certainly the most pithy) description of academic abdication is the observation "that the rules governing judicial review have no more substance at the core than a seedless grape." Ernest Gellhorn & Glen O. Robinson, *Perspectives on Administrative Law*, 75 COLUM. L. REV. 771, 780 (1975).

21. *See infra* Part II.C.

Courts and scholars have simply not devoted comparable attention to the interpretation of scope of review standards relating to facts. Instead, explanation of how these standards work has been relegated to the realm of "inarticulate" decision factors that are said to defy rules of consistency.[22] Combining these subjective factors with more objective outcomes assessments is the focus of this Article. This need not be a controversial undertaking. The scope of judicial review of facts is far less politically charged than substantive review of informal rulemaking,[23] or review of pure questions of law or policy.[24] By turning

22. The 1941 Attorney General's Report lists these "inarticulate" factors:

> In exercising their powers of review, the courts have been influenced, it is commonly thought, by a variety of inarticulate factors: The character of the administrative agency, the nature of the problems with which it deals, the nature and consequences of the administrative action, the confidence which the agency has won, the degree to which the review would interfere with the agency's functions or burden the courts, the nature of the proceedings before the administrative agency, and similar factors.

1941 ADMIN. PROC. REPORT, *supra* note 1, at 91.

23. Scope of review provisions take on a whole different dimension when they are applied to informal rulemaking. Rulemaking reversal rates in the District of Columbia Circuit, for example,

even a portion of the energy the Court expends on those decisions to scope of review over questions of fact may produce a better understood and improved scope of review doctrine.

C. Congressionally Defined Scope of Review Standards

---

range between 40-50% under narrow scope of review standards such as arbitrary and capricious, which far exceeds the reversal percentages hypothesized in this Article. *See* Patricia M. Wald, *Regulation at Risk: Are Courts Part of the Solution or Most of the Problem?*, 67 S. CAL. L. REV. 621, 636-38 (1994). Interestingly, policymaking through adjudication seems to fare better on review than policymaking through rulemaking. *See* Richard J. Pierce, Jr., *Two Problems in Administrative Law: Political Polarity on the District of Columbia Circuit and Judicial Deterrence of Agency Rulemaking*, 1988 DUKE L.J. 300, 301; *see also* Sidney A. Shapiro & Richard E. Levy, *Judicial Incentives and Indeterminacy in Substantive Review of Administrative Decisions*, 44 DUKE L.J. 1051, 1065-68 (1995) (describing the difficulties of applying arbitrary and capricious review to agency law and policy decisions).

24. *See* Chevron U.S.A., Inc. v. Natural Res. Def. Council, 467 U.S. 837 (1984); Thomas W. Merrill & Kristin E. Hickman, Chevron*'s Domain*, 89 GEO. L.J. 833, 836 (2001) (suggesting that *Chevron's* hegemony over deference to agency interpretations of law should not be so complete as to obscure the "venerable decision" of *Skidmore v. Swift & Co.*, 323 U.S. 134 (1944)); *see also* United States v. Mead, 533 U.S. 218, 221 (2001) (denying *Chevron* deference to U.S. Customs Service ruling letters on various grounds, but granting ruling letters the "eligibl[ity] to claim respect according to its persuasiveness" under *Skidmore*).

The key phrases used to define the appropriate scope of review are well known. They derive from the APA or from various agency –specific review statutes. In shorthand form, they are: arbitrary and capricious, substantial evidence, clearly erroneous, and de novo.[25] These four standards are listed in "telescopic" order, from narrow to wide scope or breadth.[26] Congress creates these standards to invite narrow to wide judicial oversight of administrative action, or, alternatively, wide to narrow deference to agency action.[27] Indeed, under de novo review, there should be no deference at all.[28]

Judicial attempts to interpret these standards accept the above progression, but they do not dictate a deference scale. These standards might be thought of as a kind of grading curve set by Congress. Under this approach, the arbitrary and capricious standard, the narrowest review, is equivalent to pass/fail,[29] a grade intended to produce a high pass rate. The substantial evidence and clearly erroneous tests, the middle standards, translate into "C" or "B" grades where most appeals

---

25. Section 706 of the APA provides that a court may

> hold unlawful and set aside agency action, findings, and conclusions found to be:
> (A) arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law; ...
> (E) unsupported by substantial evidence in a case subject to sections 556 and 557 of this title or otherwise reviewed on the record of an agency hearing provided by statute; or
> (F) unwarranted by the facts to the extent that the facts are subject to trial de novo by the reviewing court.

5 U.S.C. § 706 (2000); *see also* Veterans Judicial Review Act, 38 U.S.C. § 7261(a)(4) (2000) (requiring a clearly erroneous test for review of Board of Veterans Appeals decisions); discussion *infra* Part III.B. The clearly erroneous standard applies to judicial review of district court factual decisions as well. *See* Easley v. Cromartie, 532 U.S. 234, 237 (2001).

26. To verify this order through established case law, consult DAVIS & PIERCE, *supra* note 10, § 11.4.

27. Congress rarely articulates the reasons behind its choice of these standards, but that does not mean it is unaware of their meaning. Occasionally members of Congress show remarkable awareness of even subtle differences in oversight standards. *See infra* note 142 and accompanying text (discussing the statement of Senator Cranston concerning the choice of the clearly erroneous standard for veterans' disability review).

28. In this way, de novo review of facts is similar to the standard applied to agency decisions about law, at least as that concept was developed pre-*Chevron*. *See, e.g.*, NLRB v. Hearst Publ'ns, Inc., 322 U.S. 111 (1944).

29. *See* Thomas O. McGarity, *Some Thoughts on "Deossifying" the Rulemaking Process*, 41 DUKE L.J. 1385, 1453 (1992) (suggesting that judges reviewing complicated scientific judgments utilize a pass/fail standard to grade the agency). *But see* Richard J. Pierce, Jr., *Seven Ways to Deossify Agency Rulemaking*, 47 ADMIN. L. REV. 59, 95 (1995) (doubting whether a pass/fail standard would be enforced or followed by reviewing judges).

might be expected to result in affirmances. De novo review, the toughest standard, is more like an "A" grade where one might expect fewer cases to survive judicial scrutiny.[30]

30. The curve stipulated for use in this context approximates the (generous) grading curves at many law schools today. If one wanted to respect original intent, the grading curve in use in law schools when the APA was passed would produce much lower grades.

This grading scale is based upon a purely abstract affirmance/reversal rate. It bears no relationship to the work done by George Priest and Benjamin Klein that found success rates in litigated cases to be close to 50% regardless of whether the standard of decision was strict liability or negligence.[31] Yet the theory behind this work, that substantive standards will not independently vary outcomes because parties will have taken the applicable standard into account in deciding whether to settle or litigate, cannot be ignored.[32] The inevitability of the 50% outcome has to be considered in terms of scope of review as well.

The Priest and Klein data have limitations which could affect the 50% affirmance rate for the types of cases explored here. For one thing, gains and losses must be equal between the parties for the 50% affirmance rate to hold, which is often not the case in litigation involving the government.[33] In addition, a necessary condition for the Priest and Klein case selection hypothesis is the availability of settlement in lieu of litigation.[34] In the examples discussed here settlement is often not an option, at least not in any formal way.[35] For these reasons, it is possible to contemplate affirmance or reversal rates which depart from the 50% win/loss proposition. Hence the following chart:

Chart 1. Hypothesized Affirmance Rates
Under Various Standards of Review

Standard of Review[36]	Hypothesized Affirmance Rates
No Review	100%
Arbitrary and Capricious	85-90%

31. George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984).

32. *Id*. at 4-5 & n.17.

33. For example, in recent decades, the government's success ratio in antitrust cases, whether brought by the Department of Justice (DOJ) or the Federal Trade Commission (FTC), is 75%. Richard A. Posner, *A Statistical Study of Antitrust Enforcement*, 13 J.L. & ECON. 365, 381 (1970).

34. *See* Priest & Klein, *supra* note 31, at 12.

35. *See infra* note 103 concerning informal settlement options in Social Security Administration (SSA) disability and Freedom of Information Act (FOIA) cases.

36. "No review" can result from specific legislation, e.g., veterans' disability claims are committed to agency discretion by law. *See* 38 U.S.C. § 511(a) (2000). "Clearly erroneous" is distinguished from "substantial evidence" in theory although the two standards are often equated in practice. *See infra* notes 43, 54.

Substantial Evidence	75-85%
Clearly Erroneous	70-80%
De Novo	40-50%

Naturally, speculating on grades and setting a curve are two different exercises. Because the courts administer Congress' curve, they are free to give the grades they want. Like professors, reviewing judges sometimes think they know an "A" or an "F" when they see one, and grade accordingly. In the academic world, this tendency to subjectify the grading process can be reined in by the use of mandatory curves.[37] Such a technique is obviously not available to Congress or the courts, although some agencies with large and repetitive caseloads, such as the Social Security Administration (SSA), have on occasion experimented with the concept.[38]

---

37. Law schools impose grading curves on faculty in order to ensure some measure of grading equality in multi-sectioned courses. For example, Cardozo Law School's curve requires a 3.2 average grade in classes with more than twenty-five students. This equals a median grade slightly higher than a "B." By opting for grading equality in large sections, law schools favor consistency above accuracy, what can be called the Ted Williams hypothesis. *See* Anderson, *supra* note 9.

38. *See infra* notes 117-20 and accompanying text. To draw conclusions about the impact of scope of review standards requires a large volume of administrative decisions. Volume and

But Congress is not helpless—it can still emphasize differences among the standards.[39] It does this by implanting the suggestion that narrow review cuts in the direction of limited reversal rates and wide review cuts the other way. At this level of generality, the Supreme Court could agree that Congress has enacted an oversight spectrum, and, in the *Universal Camera* argot, "expressed a mood"[40] in which reversals of agency action should move in a more consist-ent direction.

---

repetitive fact patterns are necessary in order to develop confidence levels about the accuracy of any given curve. Even then, a reviewing court can apply a curve only when it has a significant number of cases to consider at any time. To do this, a court would have to hold decisions until it had a critical mass on which to rule, an impractical situation. Since all decision making is time sensitive, even high volume decision contexts often lack sufficient numbers at any one time to establish precise curves. *See* JERRY L. MASHAW ET AL., SOCIAL SECURITY HEARINGS AND APPEALS 139 (1978) (discussing this problem in the context of SSA disability cases).

39. Congress usually does not spend much time on scope of review matters, preferring to leave them to the default review provisions of the APA. On occasion, however, Congress has become energized about scope of review. The Bumpers Amendment, S. 1080, 97th Cong. § 5 (1981), was one such occasion. The purpose of this amendment to the APA was to encourage courts to give less deference to agency views on questions of law, but not on questions of fact. *See* Ronald M. Levin, *Review of "Jurisdictional" Issues under the Bumpers Amendment*, 1983 DUKE L.J. 355. The amendment failed to pass, of course, and shortly thereafter the Court moved in the opposite direction with the *Chevron* doctrine. *See* discussion *infra* note 54.

40. Universal Camera Corp. v. NLRB, 340 U.S. 474, 487 (1951).

Using consistent standards in this way can serve larger interests as well. In effect, they can help mediate the institutional relation-ship between agencies and reviewing courts in the fact-finding process, not unlike what the *Chevron* doctrine tries to do for those institutions in the policy-making process. By telling courts to review agency decisions on what amounts to a sliding scale, Congress can be seen as making workload choices for the agencies. Agency decisions that might pass muster under arbitrary and capricious review could be upset under a de novo standard. Consistent application of these standards of review could help guide agencies to process decisions with more or less formality. Depending on how the agency has structured its hearing process, reversals might then be expected to vary based upon the level of procedural formality chosen or by the level of explanation provided.[41] These tradeoffs rarely occur, however, because Congress and reviewing courts do not give agencies sufficiently clear signals about how the standards should operate in practice.

We are therefore left with a weak proposition. Review standards should not be directly tied to outcomes, but they should not ignore or contradict outcomes either. Scope of review standards that fail in some broad way to reflect the verbally defined sliding scale created by Congress would seem to frustrate Congress' purposes, waste agency and court time, and fail to guide the public perception of judicial review. Awareness of outcomes offers a potential feedback loop that shows how the relationship between agencies and courts is working, and it can help agencies and courts understand the effectiveness of their decision processes in relation to the goals set by Congress.

D. Unpredictability of Reversal Rates

41. Most agency adjudications are not bound by the formal adjudication provisions of §§ 554, 556, and 557 of the APA. Rather, they are informal adjudications that receive little guidance from the APA and have their procedures set by agency regulations. *See* Paul R. Verkuil, *Reflections Upon the Federal Administrative Judiciary*, 39 UCLA L. REV. 1341, 1343-44 (1992) (describing a "federal administrative judiciary" that presides over mostly informal adjudications and exceeds the federal judiciary by a factor of three); *see also* Paul R. Verkuil, *A Study of Informal Adjudication Procedures*, 43 U. CHI. L. REV. 739 (1976) (describing a study of forty-two informal adjudications).

Even if we assume reviewing courts and agencies describe review formulas in ways that invite a deference scale,[42] if not a grading curve, that does not make them predictable. Reversal rates on a statute-by-statute, or even agency-by-agency, basis are scattered across a wide range. Although outcomes may sometimes converge between close cousins like the substantial evidence and clearly erroneous tests,[43] reversal rates remain unpredictable, or even counterintuitive, between the extremes of the arbitrary and capricious and de novo standards.[44] As will be shown next, the arbitrary and capricious standard can on occasion produce more reversals than the de novo standard. In terms of the grading analogy, this is like making students do "A" work in order to get a "pass" in a pass/fail course.

---

42. In cases from *Universal Camera Corp.* to *Dickinson v. Zurko*, 527 U.S. 150 (1999), the Court has been concerned with both the unifying force of the APA scope of review provisions and the specific conclusion that the substantial evidence standard is a less stringent test than clearly erroneous review. *See* Richard J. Pierce, Jr., *Legislative Reform of Judicial Review of Agency Actions*, 44 DUKE L.J. 1110, 1114-17 (1995) (praising, with few exceptions, the consistency and the durability of the substantial evidence test).

43. The APA standards of arbitrary and capricious and substantial evidence can converge when the former is used to determine factual support in the record. *See* Ass'n of Data Processing Serv. Orgs., Inc. v. Fed. Reserve Bd., 745 F.2d 677, 684 (D.C. Cir. 1984) ("We have noted on several occasions that the distinction between the substantial evidence test and the arbitrary or capricious test is 'largely semantic.'") (citations omitted).

44. *See infra* notes 116-18 and accompanying text.

The problem seems not to be with the words themselves, but with the limited expectations for predictability they currently engender. Thus, although it is tempting to ask Congress to enhance predictability by inventing new phrases,[45] such efforts might only interject more words that require interpretation.[46] Whatever else can be said about them, the existing formulas have the virtue of familiarity.[47] The challenge is to match the existing words to a desired level of oversight rather than to ask Congress to re-conceptualize them and run the danger of further complicating the interpretive process.

II. THE SUPREME COURT BELIEVES THESE WORDS MATTER

The Supreme Court is no stranger to the application of scope of review standards. Although the cases do not speak directly in terms of outcomes, they accept the congressional deference scale from which such conclusions could be drawn. Indeed, because the Court takes scope of review and burden of proof standards seriously, the relationship between the standards employed and the outcomes obtained begs to be analyzed.

---

45. Professors Shapiro and Levy, for example, propose amendments to § 706 regarding the arbitrary and capricious standard for agency rule and policymaking that they argue could instruct the courts more explicitly. *See* Shapiro & Levy, *supra* note 23, at 1072-79. Alternatively, should Congress want closer review than arbitrary and capricious, it might adopt the "searching and careful" review standard advocated in *Overton Park*. *See supra* note 10 and accompanying text.

46. *See* Pierce, *supra* note 42, at 1132 (questioning the desirability of creating new scope of review standards along the Shapiro/Levy line because of "uncertainty over the meaning of new words").

47. Justice Jackson noted that the APA was meant to settle "long-continued and hard-fought contentions, and enact[] a formula upon which opposing social and political forces have come to rest." Wong Yang Sung v. McGrath, 339 U.S. 33, 40 (1950). That spirit should animate § 706 as well.

In Justices Breyer and Scalia, the Court has two former administrative law professors who clearly enjoy debating the meaning and role of the APA and its scope of review provisions.[48] In *Dickinson v. Zurko*,[49] Justice Scalia joined Justice Breyer's majority opinion which determined that the APA's substantial evidence standard, rather than the clearly erroneous standard, was the appropriate test when reviewing decisions of the Patent and Trademark Office.[50] This decision reversed a unanimous en banc decision of the Federal Circuit.[51] The question at issue was whether use of the clearly erroneous standard amounted to an "additional requirement[] imposed by statute or

48. Not since the days of Felix Frankfurter and Robert Jackson has the administrative law profession been as well represented on the Supreme Court. In addition to teaching administrative law, Justice Scalia was a former agency official and chairman of the Administrative Conference of the United States, and Justice Breyer was a law professor who also served as chief counsel of the Senate Judiciary Committee. Both have written important articles and books on the subject. *See, e.g.*, STEPHEN G. BREYER ET AL., ADMINISTRATIVE LAW AND REGULATORY POLICY (4th ed. 1998); Antonin Scalia, *Judicial Deference to Agency Interpretations of Law*, 1989 DUKE L.J. 511.

49. 527 U.S. 150 (1999).

50. *Id.* at 161. *Zurko* was argued on the respondent's side by another well-known administrative law professor, Ernest Gellhorn. An amicus brief urging reversal was submitted by yet another administrative law academic, John Duffy. Brief Amici Curiae of Intellectual Property Professors in Support of Petitioner, Dickinson v. Zurko, 527 U.S. 150 (1999), *available at* http://jurist.law.pitt.edu/amicus/dickinson_v_zurko.htm (last visited Nov. 18, 2000).

51. *In re* Zurko, 142 F.3d 1447 (Fed. Cir. 1998) (en banc).

otherwise recognized by law"[52] under § 559 of the APA. Application of this exception would allow the clearly erroneous standard to survive the otherwise mandatory require-ments of § 706.[53] The Court concluded, on the contrary, that the "somewhat stricter" clearly erroneous standard should give way to the unifying formula of the APA.[54]

52. 5 U.S.C. § 559 (2000).

53. *Zurko*, 527 U.S. at 153-54.

54. *Id.* at 153, 165. Subsequently, the Federal Circuit chose the APA's substantial evidence standard over the available alternative of the arbitrary and capricious standard on the presumed ground that the former standard was more strict than the latter. *See In re* Gartside, 203 F.3d 1305, 1315 (Fed. Cir. 2000). On remand in *Zurko*, the Federal Circuit applied the substantial evidence standard in reversing the Patent and Trademark Office. *See In re* Zurko, 258 F.3d 1379, 1381 (Fed. Cir. 2001).

By using the term "stricter," the Court must have assumed that reviewing courts were to look more closely and critically at cases subject to clearly erroneous review. Indeed, Justice Breyer postulated a higher reversal rate under the clearly erroneous standard when he stated: "The upshot in terms of judicial review is some practical difference in outcome depending upon which standard is used."[55] Justice Breyer, however, then pulled his punches by suggesting that the choice between these scope of review standards might not make a difference in outcome.[56] This observation must have bemused the Federal Circuit, which presumably had wanted to preserve the clearly erroneous standard precisely because it was tougher on outcomes.[57] Ultimately, by giving definitive support to the APA scope of review formulas, the Supreme Court in *Zurko* endorsed the idea of a sliding scale of reviewability with outcome consequences.

---

55. *Zurko*, 527 U.S. at 162.

56. *Id.* at 163-64; *see also* Ass'n of Data Processing Serv. Orgs. v. Fed. Reserve Bd., 745 F.2d 677, 686 (D.C. Cir. 1984) (holding that a statutory "'substantial evidence' requirement applicable" to the review before the court "demand[ed] a quantum of factual suppport no different from that demanded by the substantial evidence provision of the APA, which is in turn no different from that demanded by the arbitrary or capricious standard"); Robert L. Stern, *Review of Findings of Administrators, Judges and Juries: A Comparative Analysis*, 58 HARV. L. REV. 70, 81 (1944) (reaching the same conclusion based on pre-APA cases). The *Zurko* Court may have made this observation to reassure the Federal Circuit that things need not change under the new oversight regime. At the oral argument, a member of the Court asked counsel for Respondent (Professor Gellhorn) whether he knew of other cases in which the standard of review would produce different outcomes. He replied in the negative. *See* Oral Argument Before the Supreme Court (Mar. 24, 1999), *in* Transcript of Oral Argument, U.S. TRANS. LEXIS, at *24-*25, Dickinson v. Zurko, 527 U.S. 160 (1999) [hereinafter Gellhorn Argument].

57. The Federal Circuit's lengthy historical analysis contrasting the clearly erroneous standard with the other APA alternatives is itself evidence that scope of review matters. *See Zurko*, 142 F.3d at 1452-58. In the original *Zurko* opinion, the Federal Circuit observed: "Our ability to oversee complex legal determinations such as obviousness would be undermined if the board's underlying factual determinations were reviewed more deferentially than for clear error." *Id.* at 1459. There was little direct discussion of outcome effects except for one example where it did not matter. *See id.* at 1453-54, 1459. In fact, in a review of the reversal rates of the eighty-nine pre-APA clearly erroneous United States Court of Customs and Patent Appeals (the precursor to the Federal Circuit) patent cases cited in Justice Breyer's *Zurko* Appendix, only ten of those cases were reversed in whole or in part. *Zurko*, 527 U.S. at 165-70. This yields a reversal rate of about 11%, which tracks more closely to the hypothesized arbitrary and capricious rate than it does to the clearly erroneous rate. This reversal rate is set against the 75% grant rate. Oral argument in *Zurko* revealed that in 1998 the Patent and Trademark Office received 200,000 patent applications and granted 152,000, approximately a 75% grant rate. *See* Gellhorn Argument, *supra* note 56, at *44-*45.

Of course, the Court is not of one mind regarding the meaning of these review standards. Two years before *Zurko*, in *Allentown Mack Sales & Services, Inc. v. NLRB*,[58] Justices Scalia and Breyer wrote warring opinions on how the substantial evidence standard should be applied to a National Labor Relations Board (NLRB) decision. Justice Scalia's majority opinion utilized the APA substantial evidence standard aggressively to challenge the NLRB decision to deny union representation.[59] Justice Breyer, in dissent, believed that this approach would "weaken the system for judicial review of administrative action that [the] Court's precedents ha[d] carefully constructed over several decades."[60] "Weaken" in this context means to weaken deference, not to weaken judicial review; and that word nods in the direction of outcome analysis.

Even with their differences, *Zurko* and *Allentown Mack* demonstrate that the Court takes review standards seriously. Although neither case resorts to anything like a grading approach to deference, they both imply that scope of review can be outcome determinative. The cases also reflect the centrality of the judicial role in these matters. In *Zurko*, the Court rejected a formula established by a unanimous court of appeals and, in *Allentown Mack* the Court rejected a long-preferred NLRB approach to scope of review. Nonetheless, "if arbitrary and capricious" can mean the same thing as "clearly erroneous" in *Zurko*, and if "substantial evidence" can behave like "de novo" in *Allentown Mack*, we seem to have a Court willing both to honor and at the same time to deconstruct legislative words.

---

58. 522 U.S. 359 (1998).

59. *Id.* at 364-66.

60. *Id.* at 397 (Breyer, J., concurring in part and dissenting in part). In *Zurko*, Justice Breyer had seen no practical difference in scope of review standards because the Patent and Trademark Office has sufficient expertise to make the decision under either standard. *Zurko*, 527 U.S. at 163. In *Allentown Mack*, however, the granting of deference to the NLRB, based on the substantial evidence standard, was the issue that divided the Court. *Allentown Mack*, 522 U.S. at 388-89 (Breyer, J., concurring in part and dissenting in part).

A. Contrasting the Outcome Effect of Standards of Proof

Although the Court may not calibrate outcomes when it applies scope of review standards, it is more outcome oriented in dealing with burdens or standards of proof. In this context, the Court makes choices based on what it perceives to be the likely reversal rates of relevant formulas. The potential for outcomes to vary can be sufficient reason to invoke the Due Process Clause.

In *Santosky v. Kramer*,[61] the Court held unconstitutional a state law permitting termination of parental rights by administrative officials pursuant to a preponderance of the evidence standard.[62] The Court found that due process could only be satisfied by a clear and convincing evidence standard.[63] The *Santosky* Court was not engaged in simple statutory interpretation as was the *Zurko* Court. By rejecting legislative choice, the Court had to believe that the standard chosen made a difference in outcome; if it did not, the case should have been affirmed under the harmless error doctrine. In opting for the clear and convincing standard (as opposed to the preponderance of the evidence standard), the Court calculated that more cases would uphold parental rights, or, alternatively, that the state, as representative of the children, would lose more cases. In this way, the Court's choice of standards of proof is an outcome determinative exercise.[64]

---

61. 455 U.S. 745 (1982).

62. *Id.* at 747-48.

63. In reaching its conclusion, the Court noted that thirty-five states used the higher standard of proof. *Id.* at 749-50; *see also* Cruzan v. Dir., Mo. Dep't of Health, 497 U.S. 261, 282 (1990) (finding the clear and convincing standard in proceedings to terminate an incompetent patient's life support to meet due process concerns).

64. Unlike the Court's ambivalence towards outcomes in *Zurko*, there was no suggestion that

the application of the clear and convincing test instead of the preponderance standard might not make a difference in outcome. Instead, in comparing the two standards, the Court focused on outcome effects in the marginal case. Justice Blackmun observed: "A standard of proof that by its very terms demands consideration of the quantity, rather than the quality, of the evidence may misdirect the factfinder in the marginal case." *Santosky*, 455 U.S. at 764.

But here is the irony: Even when the Court intends to shift outcomes, it cannot ensure such results. On remand in *Santosky*, the New York state court simply upheld under the stricter standard the earlier family court decision terminating parental rights.[65] In sum, the Court's choice of differing standards of proof encourages, but does not guarantee, a different set of outcomes. The dissenters in *Santosky* did not disagree with this outcomes approach. They also phrased the inquiry in outcome terms, i.e., determining who should bear the risk of error as between the state and the individual.[66]

This test, first articulated by Justice Harlan in *In re Winship*,[67] a case determining the burden of proof to be used in criminal proceedings involving juveniles, assumes that close cases should be decided in favor of the individual rather than the government.[68] The risk of error test

---

65. *In re* John "AA," 453 N.Y.S.2d 942, 946 (N.Y. App. Div. 1982).

66. *Santosky*, 455 U.S. at 787-88 (Rehnquist, J., dissenting). Under the preponderance of the evidence standard, the risk of error is allocated evenly, and under the clear and convincing standard, it is allocated in favor of the parents. The dissent queried whether this tilt to the parents was fair to the children, a proposition the majority had considered and rejected. *Id.* at 788-91 (Rehnquist, J., dissenting).

67. 397 U.S. 358 (1970).

68. *Id.* at 370 (Harlan, J., concurring) (noting that for juveniles as well as adults, convicting the innocent is more damaging than freeing the guilty, and agreeing that beyond a reasonable doubt is the correct standard to apply in juvenile cases). Justice Brennan's opinion for the majority recognized that the outcome would have been different because the juvenile would have had to have

usefully refines the *Santosky* majority's concern as expressed in Justice Blackmun's opinion, even though the degree to which outcomes may change surely varies from context to context.[69]

---

been acquitted pursuant to the higher standard. *Id.* at 367-68.

69. The only way to know what effect shifting the result in the marginal case produces is to know how many cases cluster around the middle. That would require a study of the determinative factors in decisions. Such data is not easy to obtain when dealing with a disparate set of administrative decoders in custody cases employed by judges whose views must vary greatly. A study of the impact of personal differences was undertaken in connection with SSA disability decisions in which ALJ decisions were subject to a multifactor analysis. This study allowed researchers to utilize surveys of ALJs to determine statistically the degree to which ALJs' views about the substance of the disability programs affected their decisions. *See* MASHAW, *supra* note 38, at 19-27.

This still leaves a working hypothesis with limited predictive powers. If the burden is tilted to the government in the marginal case, we can assume individuals will record more wins than losses over time. This assumption drives the Court's intervention. In holding that juveniles are entitled to a reasonable doubt standard as opposed to the preponderance standard, the Court intended to achieve results. As Justice Harlan emphasized in his *Winship* concurrence, the choice of a standard of proof reflects "a very fun-damental assessment of the comparative social costs of erroneous factual determinations."[70] The assessment of social costs is another way of characterizing outcomes analysis.

B. Standards of Proof and Scope of Review Compared

70. *In re* Winship, 397 U.S. at 370 (Harlan, J., concurring).

Standards of proof differ from scope of review standards in two relevant ways. First, standards of proof are applied by a lower court or agency rather than by the reviewing court. Second, standards of proof are subject to constitutional review, which is rarely the case for scope of review standards.[71] Nonetheless, when Congress has spoken either on scope of review or on standards of proof, the Court tries to honor Congress' wishes. For example, in *Steadman v. SEC*,[72] the Court held that the SEC's use of a preponderance of the evidence standard in disciplinary proceedings, rather than the clear and convincing standard sought by petitioners, was compelled by § 556(d) of the APA.[73] That provision requires the support of substantial evidence in fact finding. The Court distinguished two uses of the substantial evidence standard—one for the agency under § 556(d) and one for the reviewing courts under § 706(2)(E)—and concluded that preponderance of the evidence and substantial evidence were equal standards.[74] When Congress is silent, the Court will "choose" the standard that best fits its notions of fair procedure.[75] In *Woodby v. INS*,[76] the Court held that Congress had failed to establish a standard of proof for deportation cases and rejected the civil preponderance standard in favor of the more rigorous "clear, unequivocal and convincing evidence" standard.[77]

Nevertheless, as *Steadman* reminds us, standards of proof and scope of review standards are different, and the use of a given standard at the agency level does not necessarily require its use on appeal.[78] Indeed, a reverse correlation may apply. A higher stan-dard of proof applied by an agency might be reviewed on a more deferential scope of review standard precisely because the Court and Congress are entitled to assume that a more thorough job was done by the agency.

71. *See supra* note 3.

72. 450 U.S. 91 (1981).

73. *Id.* at 102.

74. The Court cited *Vermont Yankee* in emphasizing that it was bound by Congress' choice of procedures. *Id.* at 104 (citing Vt. Yankee Nuclear Power Corp. v. Natural Res. Def. Council, 435 U.S. 519, 524 (1978)).

75. *Steadman*, 450 U.S. at 95 (citing Woodby v. INS, 385 U.S. 276, 284 (1966)).

76. 385 U.S. 276 (1966).

77. *Id.* at 277, 284-87; *see also* Vance v. Terrazas, 444 U.S. 252, 266-67 (1980) (5-4 decision) (upholding Congress' preponderance of the evidence standard in denaturalization cases). In *Woodby*, Justice Clark argued that Congress had in fact established substantial evidence as the standard of proof based on the record as a whole, which equated with the preponderance standard. *Woodby*, 385 U.S. at 287-91 (Clark, J., dissenting).

78. *Steadman*, 450 U.S. at 102.

Conversely, if the agency employs a lower standard of proof, the reviewing court might utilize a more intense standard of review, assuming Congress permits such a choice.

The *Winship* Court's cost-of-error formulation for standards of proof provides a useful way to view the impact of outcomes on agency review situations more broadly. In choosing the de novo standard rather than the arbitrary and capricious test, for example, Congress can be seen as placing the risk of error more upon the government than the individual. This more intense level of review means that individuals are more likely to be protected and agency actions are more likely to be reversed. As the *Zurko* Court indicated, the APA scope of review standards were intended to establish unifying formulas.[79] These formulas become, in essence, congres-sional methods for allocating the risk of error. Once allocated, the risk of error should affect reversal rates in the marginal case and produce more outcome uniformity. By enacting the scope of review provisions in the APA and in other statutes, Congress could be viewed as endorsing the judicial assumptions emanating from cases like *Santosky* and *Winship*.

79. Dickinson v. Zurko, 527 U.S. 150, 154 (1999); *see also* Universal Camera Corp. v. NLRB, 340 U.S. 474, 489 (1951).

The Court, however, occasionally ignores the unifying role of the APA. *Citizens to Preserve Overton Park, Inc. v. Volpe*[80] is the classic example. The Court energized arbitrary and capricious review in a way that seemed to ignore congressional choices;[81] and then the Court seemed to be taken by surprise when lower courts promptly complied.[82] Sensing the unintended consequences of its interpre-tation, the Court sought to rein in arbitrary and capricious review in subsequent decisions, rejecting any implication in *Overton Park* that trial de novo should serve as a backup to arbitrary and capricious review under the

80. 401 U.S. 402 (1971).

81. *See id.* at 419-20 (rejecting the agency affidavits on which the lower court had based its findings as "post hoc rationalizations" and remanding the case to the district court to conduct a plenary review of the agency's decision).

82. *Overton Park* quickly became the leading case on scope of review over informal adjudication and spread to review of rulemaking as well. *See, e.g.*, Ethyl Corp. v. EPA, 541 F.2d 1 (D.C. Cir. 1976), *cert. denied*, 426 U.S. 941 (1976); *see also* Peter L. Strauss, *Revisiting* Overton Park*: Political and Judicial Controls Over Administrative Actions Affecting the Community*, 39 UCLA L. REV. 1251, 1263 (1992) (calling *Overton Park* the "foundation stone for contemporary 'hard look' judicial review").

APA.[83] Although the Court did not rein in "hard look" arbitrary and capricious review in other settings,[84] it minimized intense review of informal agency adjudications by district courts in an effort to avoid the potentially higher reversal rates such review surely would have produced.

83. In *Camp v. Pitts*, 411 U.S. 138 (1973) (per curiam), the Court reversed the Fourth Circuit's decision under the arbitrary and capricious standard to subject the Comptroller of the Currency's denial of a new bank application to district court trial review. *Id.* at 142. In *Pitts v. Camp*, the lower court had resorted to de novo review because the Comptroller twice failed to explain its decision properly. 463 F.2d 632, 634 (4th Cir. 1972). In the Fourth Circuit's view, *Overton Park* compelled de novo review when the procedures employed were inadequate to support the agency decision. *Id.*

84. Hard look review has retained vitality in connection with arbitrary and capricious review of informal rulemaking. *See* Greater Boston Television Corp. v. FCC, 444 F.2d 841, 851-52 (D.C. Cir. 1970). But because of *Camp v. Pitts*, hard look review has not gained much of a foothold in informal adjudication review. *See* Baltimore Gas & Elec. Co. v. Natural Res. Def. Council, 462 U.S. 87, 101-06 (1983) (criticizing the "hard look" dimension of the arbitrary and capricious standard implied in *Overton Park*); Paul R. Verkuil, *Judicial Review of Informal Rulemaking: Waiting for* Vermont Yankee II, 55 TUL. L. REV. 418, 424 (1981) (characterizing the *Overton Park* approach as "intrusive substantive review"); *see also* Pierce, *supra* note 23, at 308-13 (discussing the "powerful deterrents" courts have imposed on agency rulemaking); Wald, *supra* note 23, at 625-29 (discussing the "hard look" doctrine in rulemaking).

In sum, when it comes to scope of review, the Court tries to honor legislative expectations.[85] By contrast, in standard of proof situations, the Court questions legislative choice especially at the state level and occasionally plays a trump card—the Due Process Clause. In either situation outcomes are relevant to understanding both the differences among the various review formulas and the Court's guidance to lower courts.[86]

---

85. *See* Consolo v. Fed. Mar. Comm'n, 383 U.S. 607, 620 (1966) ("Congress was very deliberate in adopting [the substantial evidence] standard of review. It frees the reviewing courts of the time-consuming and difficult task of weighing the evidence, it gives proper respect to the expertise of the administrative tribunal and it helps promote the uniform application of the statute.") (footnotes omitted). Of course, when it comes to due process review, state legislatures have traditionally been treated with less deference than Congress. *Compare* Goldberg v. Kelly, 397 U.S. 254, 266 (1970) (mandating state welfare pretermination procedures), *with* Mathews v. Eldridge 424 U.S. 319, 349 (1976) (concluding that due process can be satisfied absent a hearing in federal disability cases).

86. Under the *Chevron* doctrine, the courts of appeals have bought into the Supreme Court's approach. *See* Peter H. Schuck & E. Donald Elliott, *To the* Chevron *Station: An Empirical Study of Federal Administrative Law*, 1990 DUKE L.J. 984, 1029-31 (analyzing 2,000 decisions by the courts of appeals that document a pre-*Chevron* affirmance rate of 71% versus a post-*Chevron* rate of 81%). This 10% "affirmance premium" is an indicator of lower court compliance with the instructions of the Supreme Court. It is far less clear whether *Chevron* has had the effect of ensuring deterrence by the Supreme Court itself. *See* Whitman v. Am. Trucking Ass'n, Inc., 531 U.S. 457, 481-86 (2001) (Scalia, J.) (indicating that the case represented a rare rejection of an agency's interpretation of a statute in *Chevron* Step 2); INS v. Cardoza-Fonseca, 480 U.S. 421, 453 (1987) (Scalia, J., concurring) (granting no deference to an agency's interpretation of a statute which is

C. Scope of Review in Other Settings

inconsistent with the contested phrase's plain meaning); *see also* Thomas W. Merrill, *Judicial Deference to Executive Precedent*, 101 YALE L.J. 969, 998 (1992) (stating that the Court frequently ignores *Chevron*); Thomas W. Merrill, *Textualism and the Future of the* Chevron *Doctrine*, 72 WASH. U. L.Q. 351, 354 (1994) (same).

To complete the picture it should be remembered that the Court frequently monitors scope of review situations outside the administrative context. Two disparate examples, sentencing review and punitive damages review, serve to make the point. The Court carefully monitors deference by the lower courts to the United States Sentencing Guidelines promulgated by the United States Sentencing Commission.[87] In *Koon v. United States*,[88] the Court held that courts of appeals must review district court departures from the Sentencing Guidelines under the abuse of discretion standard rather than the de novo test.[89] When it comes to questions of law, the Court reasoned that little turned on those widely disparate labels since "[a] district court by definition abuses its discretion when it makes an error of law."[90]

In *Cooper Industries, Inc. v. Leatherman Tool Group, Inc.*,[91] however, the Court held that de novo review is appropriate when reviewing the constitutional sufficiency of punitive damages awards.[92] The Court distinguished *Koon* as applying to legislatively enacted guidelines and not issues of constitutional dimension.[93] Significantly, the Court adopted an outcomes-based analysis in determining the standard of review: "[I]t does seem likely that in this case a thorough, independent review of the District Court's rejection of petitioner's due process objections to the punitive damages award might well have led the Court of Appeals to reach a different result."[94]

Even though these situations often deal with review of legal rather than factual questions, they testify to the Court's ongoing interest in the relationship of standards to outcomes. In order to carry the burden of this Article, however, we must now shift the Court's interest in outcomes from review of questions of law and policy to review over

87. *See* U.S. SENTENCING COMM., GUIDELINES MANUAL (1991).

88. 518 U.S. 81 (1996) (reviewing the federal conviction of police officers in the Rodney King case).

89. *Id.* at 96-100; *see also* Cooter & Gell v. Hartmarx Corp., 496 U.S. 384, 405 (1990) (applying the abuse of discretion standard to Rule 11 sanctions).

90. *Koon*, 518 U.S. at 100.

91. 532 U.S. 424 (2001).

92. *Id.* at 435-36; *see also* BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 568-74 (1996) (applying the Due Process Clause of the Fourteenth Amendment to excessive punitive damages awards).

93. *Cooper Indus.*, 532 U.S. at 432-33.

94. *Id.* at 441. Justice Ginsberg, in dissent, doubted the effectiveness of the new standard and was also concerned about the diminishing effect on the Seventh Amendment of setting aside a jury verdict on punitive damages. *Id.* at 444-47 (Ginsberg, J., dissenting).

questions of fact, or mixed questions of fact and law. The next Part seeks to do just that.

III. SCOPE OF REVIEW IN CONTEXT: SOCIAL SECURITY DISABILITY, VETERANS DISABILITY, AND THE FREEDOM OF INFORMATION ACT

It is not hard to find individual scope of review situations to support the reversal rate percentages postulated here; nor is it hard to find examples that contradict them. However, it is hard to find administrative adjudications in numbers sufficient to support meaningful conclusions about the percentages involved. The number of cases not only offers some level of statistical confidence, but also facilitates assumptions about the issues raised. A large number of repetitive fact situations by definition will yield cases that turn more on scope of review over facts, pure or mixed.[95]

In the modern administrative state's two largest decision systems — disability decisions by the SSA and Freedom of Information (FOIA) requests certified by the Department of Justice (DOJ)—the number of underlying claims filed annually is comparable. Each agency faces about two million requests for either benefits or documents annually. These agencies also have another roughly comparable statistic—each grants approximately 50% of the claims before it prior to judicial review.[96] In addition to these agencies, the third highest decision volume agency—the Veterans Administration disability program—will also be evaluated.

Three scope of review situations are analyzed here: (1) district court review of SSA disability cases under the substantial evidence standard; (2) Court of Appeals for Veterans Claims (CVA) review of Board of

95. "Mixed" questions of law and fact are entitled to deference similar to "pure" fact questions under the appropriate scope of review standards. *See, e.g.*, NLRB v. Hearst Publ'ns, Inc., 322 U.S. 111, 130-31 (1944). Another advantage of utilizing highly fact-based decisions is that one can assume they will, in the aggregate, turn on factual components. It would be extremely difficult and perhaps infeasible to collect case statistics in such a way as to distinguish between decisions that turn on facts, rather than law.

96. *See* SOCIAL SECURITY ADVISORY BOARD, DISABILITY DECISION MAKING: DATA AND MATERIALS 20 (2001) [hereinafter SSAB, DISABILITY DECISION MAKING]; USDOJ, FOIA POST, SUMMARY OF ANNUAL FOIA REPORTS FOR FISCAL YEAR 1999, *available at* http://www.usdoj.gov/oip/foiapost/2001foiapost17.htm. (last visited Nov. 19, 2002). The 50% grant estimate is also a reminder of the Priest and Klein selection hypothesis. *See* Priest & Klein, *supra* note 31, at 1.

Veterans Appeals (BVA) decisions under the clearly erroneous standard; and (3) district court review of FOIA claims decided under the de novo standard.[97] In the last Part, the conclusions drawn from the examples will be tested against comparable statistics drawn from appellate review of sentencing decisions.[98]

A. Social Security Administration Disability Cases

97. Of the four formulations described earlier, only arbitrary and capricious review is excluded. Since, as to review of facts, that standard is equated with the substantial evidence standard, see discussion *supra* note 43, its omission should not restrict the analysis.

98. *See infra* Part IV.

The SSA disability process encompasses major federal programs[99] that affect millions of Americans and expend close to $100 billion annually.[100] The hearing process is prolonged and complicated.[101] After the initial and reconsideration stages, which are decided on applicants' documentary submissions to state agencies, claimants get a face-to-face hearing before an Administrative Law Judge (ALJ); and, after appeals council review,[102] a federal district court may review the decision under the substantial evidence standard.

---

99. Disability is sought under either Title II of the Social Security Act, 42 U.S.C. §§ 401-33 (2000), or under Title XVI of the Act, 42 U.S.C. §§ 1381-1383d (2000), depending upon whether the claimant has worked the required number of quarters to qualify for Social Security. Sometimes claims are filed under both titles. For purposes here, both statutes will be referred to as "SSA disability" claims since the substantive standards of disability and the scope of review standards are the same.

100. The disability program currently covers about 10 million individuals who received $90 billion in support, or 5% of all federal spending in the 2001 fiscal year. About $5 billion of SSA's administrative budget is spent on disability work. *See* SOCIAL SECURITY ADVISORY BOARD, CHARTING THE FUTURE OF SOCIAL SECURITY'S DISABILITY PROGRAMS: THE NEED FOR FUNDAMENTAL CHANGE 1 (2001), *available at* http://www.ssab.gov/disabilitywhitepap.pdf (last modified Jan. 2001) [hereinafter SSAB, CHARTING THE FUTURE].

101. *See* SSAB, DISABILITY DECISION MAKING, *supra* note 96, at 4-5 (describing SSA's disability applications and appeals process).

102. Appeals Council review must be sought by claimants as a condition for judicial review. *See* 20 C.F.R. § 404.900(b) (2002). Review is discretionary with the Appeals Council, which also may trigger review without request from a claimant. *See* Charles H. Koch, Jr. & David A. Koplow, *The Fourth Bite at the Apple: A Study of the Operation and Utility of the Social Security Administration's*

Appeals Council, 17 FLA. ST. U. L. REV. 199, 243-49 (1990); *see also* Sims v. Apfel, 530 U.S. 103, 103 (2000) (refusing to apply issue exhaustion to applications before the Appeals Council).

The hearing before the ALJ is conducted de novo, on a non-adversarial basis.[103] SSA-ALJs decide more than 500,000 cases per year, or an average of thirty-eight cases per ALJ per month.[104] The SSA employs about 1100 federal ALJs who preside over these administrative hearings.[105] To put this number in perspective, SSA-ALJs are equal in number to the entire federal district and circuit court judiciary.[106]

ALJ disability hearings, conducted on a de novo basis, yield an average reversal rate in excess of 50%.[107] Because this is the first chance the claimant has to tell her story face-to-face, such a statistic seems to make sense. State officials can be expected to kick close cases to the ALJs, and many cases fall into this category.[108] Moreover, even though inter-judge reversal rates can vary greatly,[109] the overall ALJ

---

103. *See* Richardson v. Perales, 402 U.S. 389, 400-01 (1971). The adjudicatory system is termed "informal" by the SSA, which "conduct[s] the administrative review process in an informal, nonadversary manner." 20 C.F.R. § 404.900(b) (2002). Although the Supreme Court has noted that "Social Security proceedings are inquisitorial rather than adversarial," *Sims*, 530 U.S. at 110-11, the presence of ALJs as presiding officers makes disability cases functionally like formal proceedings even if not technically so under the APA. *See* Paul R. Verkuil et al., *The Federal Administrative Judiciary*, *in* 2 ADMINISTRATIVE CONFERENCE OF THE UNITED STATES: RECOMMENDATIONS AND REPORTS 779, 815-17 (1992). Disability cases are not governed by §§ 554, 556, and 557 of the APA.

104. *See* SSA ANNUAL STATISTICAL SUPPLEMENT, tbl. 2.F9 (2000) [hereinafter 2000 SSA STAT. SUPP.] (estimating 596,999 hearing decisions in 1999 and 584,546 in 2000).

105. The SSA employed approximately 1180 ALJs in 1998; approximately 1107 in 1999; and an estimated 1403 in 2000. *Id.* at tbl. 2.F8.

106. *See* Judith Resnik, *Trial as Error, Jurisdiction as Injury: Transforming the Meaning of Article III*, 113 HARV. L. REV. 924, 951-52 (2000) (listing 270 court of appeals and 850 district judges for a total of 1,120).

107. In the ten year period from 1991 through 2000, ALJ dispositions favorable to the claimant averaged between 52.9% (1998) and 67.1% (1992). SSA OFFICE OF HEARINGS AND APPEALS KEY WORKLOAD INDICATORS (Fiscal 2000), at 2 [hereinafter KEY WORKLOAD INDICATORS 2000]; *see also* MASHAW, *supra* note 38, at 21-24 (discussing the variance of reversal rates of individual ALJs); Richard E. Levy, *Social Security Disability Determinations: Recommendations for Reform*, 1990 B.Y.U. L. REV. 461, 482-84 (linking the increase in disability terminations to the SSA "caseload crisis" of the 1980s).

108. During the period between 1990 and 2000, state agency initial allowance rates moved between 20 and 30%, and reconsideration rates were between 10 and 20%. SSAB, DISABILITY DECISION MAKING, *supra* note 96, at 21. Given that the state officials are acting on documentary records only (largely medical reports and vocation information), and that only denials, but not grants, may be appealed, it makes institutional sense that all but unequivocally clear cases be denied at this stage. *See* MASHAW, *supra* note 38, at 54-57.

109. *See* JERRY L. MASHAW, BUREAUCRATIC JUSTICE: MANAGING SOCIAL SECURITY DISABILITY CLAIMS 85-88 (1983) (suggesting inter-ALJ disparities can be reduced or eliminated by imposing a grading system); *see also* Heckler v. Campbell, 461 U.S. 458, 467 (1983) (approving an agency

reversal rate falls within the range for de novo reversals postulated earlier. What happens next, however, does not.

rule that was designed to limit variances among ALJs in the application of vocational criteria). The larger problem of inserting an independent ALJ decision system within a largely bureaucratic system is not addressed by these statistics. *See* MASHAW, *supra* note 38, at 31-33. Nevertheless, because ALJs have established their independence within that system, the SSA has accepted their performance at the 50% mean. *See* Nash v. Califano, 613 F.2d 10, 12-13 (2d Cir. 1980) (reviewing the effect of the SSA's "Quality Assurance Program" on the independence of ALJs).

Even after high-side reversal rates at the agency level and a further review by the Appeals Council,[110] the reversal or remand rates that occur on the district court's substantial evidence review are unusually high. The rate at which district courts reverse and remand disability determinations exceeds 50%.[111] Between 1994 and 1998, SSA civil

110. In 1998, Appeals Council dispositions totaled 101,877 versus 618,578 total ALJ dispositions, or about 16%. SOCIAL SECURITY ADMINISTRATION ANNUAL STATISTICAL SUPPLEMENT 135 (1999) [hereinafter 1999 SSA STAT. SUPP.]. In 1999, the Appeals Council percentage was about 13% and in 2000 it grew to about 23%. *See* 2000 SSA STAT. SUPP., *supra* note 104, at 117. In 2000, the Appeals Council granted about 10% of the cases seeking reversal of ALJ denials of benefits and also remanded about 20%. *See* KEY WORKLOAD INDICATORS 2000, *supra* note 107, at 43.

111. Since 1995, remands have become an increasing part of the district court decision process. SSA now estimates that 40% of all judicial decisions result in remands and that reversals have abated accordingly. Discussion with Rita Beier, Head of the Bureau of Hearings and Appeals (July 2, 2001). In fiscal year 2000, 48% of the 12,001 district court decisions resulted in remands. *See* SSAB, CHARTING THE FUTURE, *supra* note 100, at 8. In general, remands result in grants to

litigations totaled between 8500 and 15,000 cases per year which constituted a significant portion of the federal caseload.[112] If district court remands are partially equated with reversals in terms of outcome effect,[113] in virtually every year studied the nonaffirmance total meets or exceeds expectations under the postulated de novo standard, let alone under the substantial evidence standard.[114] Indeed, about the only standard the 50% reversal rate seems to meet is Priest's case selection hypothesis.[115]

---

claimants in the vast majority of cases. *See, e.g.*, SOCIAL SECURITY ADMINISTRATION, STATISTICAL SUPPLEMENT, at tbl. 2, F10 (1995) [hereinafter 1995 SSA STAT. SUPP.] (stating that in 1999, 83% of remanded cases resulted in agency reversals; in 1993, the rate of reversal after remand was 94%).

112. *See., e.g.*, 1995 SSA STAT. SUPP., *supra* note 111, at tbl. 2.F10. Table 2.F10 of each annual supplement contains this data, which excludes remands.

113. Remands are like reversals in that they produce a high percentage of decisions in favor of claimants once they are returned to the SSA. More than 70% of judicial remands result in disability grants by the SSA. *See* discussion *supra* note 111.

114. Professors Koch and Koplow report statistics on district court reversal rates of "20% in 1982, 30% in 1983, 57% in 1984, 46% in 1985, and 38% in 1986." Koch & Koplow, *supra* note 102, at 226. In the last few years, the combined reversal/remand rate has exceeded 50%. *See supra* note 111 and accompanying text.

115. *See supra* notes 31-34 and accompanying text. The selection hypothesis relies on assumptions about rational choices by plaintiffs and defendants that may not apply in the SSA disability situation. But as the number of claimants represented by counsel increases, and informed choices are made about which cases to select, that hypothesis may become more potent. *See infra* note 229 and accompanying text.

Why are SSA disability determinations so prone to rejection in district courts even after ALJs have seemingly done their job under the de novo standard? There are no easy answers. Such a result appears to contradict the assumption that careful administrative decision making, with an intense de novo review standard, should lead to more relaxed judicial oversight.[116] Moreover, the SSA has been trying to improve the system for years, oftentimes with explicit congressional assistance,[117] yet district judges, rather than deferring under the substantial evidence test, still expend their valuable time, or that of their magistrates, in close review of the facts supporting individual cases.

116. *See supra* text accompanying note 41.

117. Congress is of two minds about the SSA disability system, wanting both to control the costs of a multi-billion dollar program and also to care for individual claimants through constituent services. *See* MASHAW, *supra* note 109, at 53-54; *see also* SSAB, CHARTING THE FUTURE, *supra* note 100, at 1 (citing the need for "in-depth review" of disability programs).

These high rates of rejection, counterintuitive from an outcomes analysis perspective, can only be explained by entering the realm of inarticulate factors. SSA disability decisions suffer from an entrenched judicial skepticism about their fairness and accuracy. To some extent, this is a function of the difficulty of applying the legal standard for determining disability.[118] Taken literally, an applicant who could perform any job in the national economy, even if that job is unavailable, would have no entitlement to disability benefits. Such a theoretical prospect is hard to accept,[119] and district courts intuitively may assist claimants by ameliorating the standard's impact.[120] Even

118. 42 U.S.C. § 423(d)(1)(A) (2000) defines disability in terms of the effect of physical and mental impairment on the person's ability "to engage in any substantial gainful activity" in the national economy. *See* Heckler v. Campbell, 461 U.S. 458, 461-62 (1983) (establishing a grid system for determining what jobs exist in the national economy). In determining disability, the SSA must consider the claimant's "residual functional capacity," age, education, and past work experience. 20 C.F.R. § 404.1520(f) (2002).

119. *See* Lance Liebman, *The Definition of Disability in Social Security and Supplemental Security Income: Drawing the Bounds of Social Welfare Estates*, 89 HARV. L. REV. 833, 843-47 (1976).

120. *See id.* at 844-45.

members of Congress who complain about the costs of reversals often intervene with "status inquiries" to the agency on behalf of their constituents who are applying for benefits,[121] a practice that may also have a subtle effect on outcomes. People without jobs and with poor medical records simply present appealing situations which may make some district courts partners in validation, rather than skeptical reviewers. In this setting, the *Winship* proposition about assigning risks of error or identifying social costs gets shifted to the government in practice, even though the scope of review standard places it upon the claimant in theory.[122]

---

121. *See The Regulatory Morass at the Centers for Medicare and Medicaid Services: A Prescription for Bad Medicine; Hearing Before the House Comm. on Small Business*, 107th Cong. (2001) (prepared remarks of Brian Seeley, Board of Directors, Power Mobility Coalition; President and CEO, Seeley Medical, Inc.) (complaining about the cost of appeals and the 80% reversal rate in Medicare appeals); COMM. ON STANDARDS OF OFFICIAL CONDUCT, OFFICE OF ADVICE & EDUC., HIGHLIGHTS OF HOUSE ETHICS RULES (2001) (summarizing the rules for members making status inquiries into agency and court cases), *available at* http://www.house.gov/ethics/Highlights2001.htm (last modified Jan. 2001). Although Congress seeks to lower the costs of reversals, the House Ethics Rules allow members to request such reversals of agency decisions. According to the Rules:

> Members have broad discretion in helping constituents: may make a status inquiry; urge prompt and fair consideration; ask for full and fair consideration consistent with applicable law and regulations; arrange appointments—or, where appropriate, express judgment, or ask for reconsideration of decision if it is unsupported by law.

Id.

122. *See supra* notes 67-70 and accompanying text.

To some extent the SSA has been its own worst enemy. During the 1980s, the SSA compromised the validity of its decision process in two ways: (1) by commencing a policy of nonacquiescence in federal court decisions;[123] and (2) by insisting on fixed percentages of affirmances and reversals through so-called Bellmon review.[124] These actions were the result of a variety of pressures. On the one hand, the SSA desired to make its decision process more consistent and predictable; at the same time, it was under pressure from Congress and the White House to reduce the rolls of disability recipients.[125] The SSA's actions alienated virtually all sides and in the process caused the courts to be skeptical about the accuracy of individual agency decisions.[126] This led to dramatically higher reversal rates in the mid-1980s.[127] Once the SSA backed off on Bellmon review,[128] SSA reversal rates began to decline.[129] None-theless, with remands now on the rise, the current nonaffirmance rate continues to exceed the percentage postulated for substantial evidence review.[130]

---

123. "Nonacquiescence" refers to the SSA's practice of limiting the precedential effect of courts of appeals' decisions to the cases before them, or of selecting precedents it desires to follow. *See* Carolyn A. Kubitschek, *A Re-evaluation of* Mathews v. Eldridge *in Light of Administrative Shortcomings and Social Security Nonacquiescence*, 31 ARIZ. L. REV. 53, 53-54 (1989) (asserting that this policy undermines the due process calculus of *Mathews*).

124. The Bellmon Amendment, Pub. L. No. 96-265, § 304, 94 Stat. 441, 456 (1980), established a performance review program in the Appeals Council that targeted ALJs with high percentages of grants. *See* Nash v. Bowen, 869 F.2d 675, 678-80 (2d Cir. 1989). For an overview of the development of nonacquiescence, see generally MARTHA DERTHICK, AGENCY UNDER STRESS: THE SOCIAL SECURITY ADMINISTRATION IN AMERICAN GOVERNMENT 139-51 (1990).

125. *See, e.g.*, Holst v. Bowen, 637 F. Supp. 145, 147 (E.D. Wash. 1986) (referring to an "anarchical situation" in which the SSA systematically ignored the Ninth Circuit); Ass'n of Admin. Law Judges, Inc. v. Heckler, 594 F. Supp. 1132 (D.D.C. 1984) (holding that the Bellmon review program violated ALJ decisional independence).

126. *See* Levy, *supra* note 107, at 506; *see also* Schweiker v. Chilicky, 487 U.S. 412, 416 (1988) (stating that the SSA conceded benefits were erroneously terminated for approximately 200,000 recipients).

127. *See* Koch & Koplow, *supra* note 102, at 226.

128. *See* Robert Pear, *Reagan Suspends Benefits Cutoff*, N.Y. TIMES, Apr. 14, 1984, at A1; *see also* Levy, *supra* note 107, at 506-07 (arguing that Bellmon programs made the federal courts less deferential to the SSA).

129. The Bellmon review process took place during the 1980s and led to some high reversal rate years. *See* Koch & Koplow, *supra* note 102, at 226 (referring to reversal rates of 46% in 1985 and 38% in 1986). Since then, reversal rates have been lower and more consistent on an annual basis. *See supra* notes 107-15 and accompanying text.

130. Remands are usually based on some procedural failure with the ALJ or agency decision and can have a disruptive effect on the SSA because the cases must be reassigned and retried or disability grants paid (the predominate outcome). *See* MASHAW, *supra* note 38, at 130.

These and other factors[131] combine effectively to override the congressionally chosen scope of review standard and raise questions about whether the SSA disability system should be in the federal courts at all, or at least in district courts on a substantial evidence basis.[132] The recently created Court of Veterans Appeals[133] offers a competing model for administrative review that may be looked to as a source of procedural reform for the SSA disability system.[134]

B. Veterans Disability Claims

Veterans disability payments, like SSA disability payments, are awarded to millions at the cost of billions.[135] There is no more favored class of beneficiaries than veterans, yet, unlike Social Security claimants, claims by veterans were judicially *unreviewable* by statute

131. Another difficult factor to measure is that, due to their high workloads, ALJs who hear cases do not write their own opinions, which leads to judicial skepticism about the quality of the underlying decision. *See* SSAB, CHARTING THE FUTURE, *supra* note 100, at 4. Moreover, under the scope of review standard, the court may remand for "good cause." 42 U.S.C. § 405(g) (2000). This provision may have the effect of encouraging remands by the district courts.

132. *See* Levy, *supra* note 107, at 528-32 (advocating "[a]n Article I Court of Disability Appeals with jurisdiction to review ALJ disability determinations" on questions of law).

133. The Veterans Administration's (VA) system of review may have its own problems. *See* James T. O'Reilly, *Burying Caesar: Replacement of the Veterans Appeals Process is Needed to Provide Fairness to Claimants*, 53 ADMIN. L. REV. 223, 243-47 (2001) (arguing for combining the VA disability appeals process into the existing SSA disability process).

134. The Judicial Conference of the United States has called for administrative review of SSA disability claims, or, alternatively, for limited judicial review after the district court phase. Recommendation 9a provides: "Legislation should be requested to improve the adjudicative process for Social Security disability claims by establishing a new mechanism for administrative review of ALJ decisions and limiting the scope of appellate review in the Article III courts." JUDICIAL CONFERENCE OF THE UNITED STATES, LONG RANGE PLAN FOR THE FEDERAL COURTS 33 (1995). The study also contemplates substantial evidence review in the district courts only, with court of appeals review generally limited to questions of law. *Id.* at 46-47.

135. *See* U.S. DEP'T OF VETERANS AFFAIRS FACT SHEET, *VA Disability Compensation Claims Processing*, at 2 (May 2001), *available at* http://www.va.gov/pressrel/claimpro.htm (stating that 2.3 million veterans received benefits in 2001 and that the 2000 fiscal year cost of these benefits was $14.7 billion); U.S. DEP'T OF VETERANS AFFAIRS ANNUAL BENEFIT REPORT 80-99 (2001), *available at* http://www.vba.va.gov/reports.htm; Press Release, U.S. Dep't of Veterans Affairs, VA Benefits Assist Millions of Americans (Sept. 12, 2000), *at* http://www.va.gov/OPA/pressrel/PressArtInternet.cfm?id=213 ("[A]pproximately 2.3 million of the nation's 24.4 million veterans [are] directly compensated each month for injuries or illnesses .... The 70-year projection shows that spending for disability compensation and survivors payments will peak at $37.3 billion in 2032 ....").

until relatively recently.[136] Under prior practice, three-person Veterans Affairs (VA) rating boards made decisions that were subject to final review by the agency. Now, judicial review of Board of Veterans Appeals (BVA) administrative decisions is conducted by the Article I Court of Veterans Appeals (CVA). Review is based on a clearly erroneous review of material facts.[137] In addition, the CVA must also apply a further (and unique) standard of proof: by statute the veteran is entitled to the benefit of the doubt.[138]

136. For a discussion of the congressional shift in 1988 from judicial unreviewability of veterans claims to reviewability, see Bill Russo, *Ten Years After the Battle for Veterans Judicial Review: An Assessment*, FED. LAW., June 1999, at 26; O'Reilly, *supra* note 133, at 223.

137. 38 U.S.C. § 7261(a)(4) (2000). The Court of Appeals for the Federal Circuit is granted the authority to review the CVA on legal errors only. The CVA has interpreted the clearly erroneous test as if it were being applied in a court review situation. *See* Gilbert v. Derwinski, 1 Vet. App. 49, 51-52 (amended 1991).

138. 38 U.S.C. § 5107(b) (2000) provides: "When there is an approximate balance of positive and negative evidence regarding any issue material to determination of a matter, the Secretary [of Veterans Affairs] shall give the benefit of the doubt to the claimant." *Id.*

This latter provision, which codified VA practice going back to the post-Civil War period,[139] complicates the application of a clearly erroneous scope of review by the CVA. In effect, the reviewing court is being asked to review decisions twice: first under the clearly erroneous standard and then under the benefit of the doubt standard. This led the CVA to adopt the view, guided by cases like *Winship* and *Santosky*,[140] that benefit of the doubt only comes in when the evidence is at "equipoise."[141] In practice, however, the presence of the "tie breaker" probably intensifies the oversight role of clearly erroneous review in every case.[142]

---

139. *See* 38 C.F.R. § 3.102 (2001) (codifying "reasonable doubt" policy); *Gilbert*, 1 Vet. App. at 55.

140. *See supra* notes 61-70 and accompanying text. In *Gilbert*, the CVA applied *Winship* and *Santosky* when it interpreted the benefit of the doubt test. *Gilbert*, 1 Vet. App. at 53-54.

141. *Gilbert*, 1 Vet. App. at 54. The CVA accepted the VA's use of a baseball analogy to describe the benefit of the doubt standard, i.e., the "tie goes to the runner." *Id.* at 55-56.

142. The clearly erroneous test itself was carefully chosen by Congress when it created the Veterans Judicial Review Act. One of its sponsors, Senator Cranston, stated that he was

> confident that utilization of the "clearly erroneous" standard ... will permit that court

Still, even when the benefit of doubt standard is applied simultaneously with the clearly erroneous test, this "double test" produces a BVA reversal/remand rate that approximates the hypothesized rate for clearly erroneous reversals.[143] Of course, the BVA's reversal rate varies over time,[144] but at no time does it meet or exceed the reversal/remand rate by district courts over SSA cases.

> to carry out a more complete analysis of factual matters than would be appropriate under an "arbitrary and capricious" standard ... It is also possible ... that the "clearly erroneous" standard could result in more fact review than if the "substantial evidence" standard ... were adopted here.

134 CONG. REC. 31,472 (1988) (statement of Sen. Cranston). This remarkably nuanced statement reflects a sophisticated congressional recognition of the differences among the three scope of review standards discussed here and lends support to the outcomes analysis approach.

143. *See Claims Processing Testimony Before the House Comm. on Veterans Affairs, Subcomm. on Oversight and Investigations*, 106th Cong. 2-3 (2000) (statement of Rick Surratt, Dep. Nat'l Legis. Dir., Disabled Am. Veterans) [hereinafter Surratt, *Claims Processing Testimony*] (stating that the BVA reversed about 18% of cases between 1992 and 1999 and that before the 1988 legislative changes, remands and reversals totaled 12% of cases); Russo, *supra* note 136, at 28 (calculating a 12-14% BVA reversal prior to judicial review and a 15-20% rate after).

144. Surratt, *Claims Processing Testimony*, *supra* note 143, at 2-3, 8; DEP'T OF VETERANS AFFAIRS, OFFICE OF INSPECTOR GENERAL AUDIT REP. # 5D2-B01-013, at 5-6 (1995), *available at* http://www.va.gov/oig152/reports/1995/5D2-B01-013%20--%20appeals.htm (citing a remand rate of 18 to 23% before the Court of Veterans Appeals was formed and a 50% remand rate after); Russo, *supra* note 136, at 28 (placing the remand rate at "over 40%.").

Although the VA review situation is closer to predictable limits, some critics of the VA claims system are unimpressed with the quality of BVA review by the CVA. As a result, two dramatically opposed suggestions have emerged: (1) that the VA system be folded into the SSA system and district judges do both jobs;[145] or (2) that the VA Article I disability court concept be expanded to encompass both VA and SSA claims.[146] These alternatives remain at the heart of reform issues in disability review.[147]

C. Freedom of Information Act Cases

FOIA provides for de novo review in district courts of agency actions denying access to documents in the possession of government.[148] Unlike district court review of SSA disability cases, the APA preferred standards of "arbitrary and capricious" and "substantial evidence" have been explicitly preempted in FOIA cases by the de novo standard.[149] In FOIA cases, the issue before the district court is usually whether the requested documents meet the specific exemptions claimed by the non-producing agency. Unlike the SSA and VA disability situations, the requester in FOIA cases has not had a hearing before the agency, since a hearing would require either production of the very documents the government is suppressing or an administrative in camera review proceeding.

What the district court sees are conflicting affidavits rather than a transcript of a hearing below, and these cases are often decided by cross motions for summary judgment.[150] Although summary judgment is by definition concerned with issues that are legal and not factual in

145. O'Reilly, *supra* note 133, at 243-47.

146. Levy, *supra* note 107, at 528-32.

147. *See infra* note 228.

148. 5 U.S.C. § 552(a)(4)(B) (2000).

149. The de novo review standard made explicit in FOIA is different from the de novo standard defined under APA § 706(2)(F). That standard only comes into play in the rare circumstances where there is no record below. *See, e.g.*, Camp v. Pitts, 411 U.S. 138, 141-42 (1973) (per curiam); *supra* note 83 and accompanying text. In FOIA cases, although there is no administrative hearing in the traditional sense, there is still a record below which gets certified to the court by the DOJ. Hence, the FOIA de novo standard allows the district court to review on the existing record.

150. It is estimated that summary judgment occurs in about 90% of FOIA appeals. Discussion with David Vladeck, Public Citizen Litigation Group (July 11, 2001).

nature, the standard of review does not change because de novo applies in both situations.[151]

151. FOIA review involves the application of scope of review to facts and/or to application of facts to law. Pure questions of law and policy are not implicated and *Chevron* deference does not apply. Since FOIA is an APA statute, it is not organic agency legislation which would normally be eligible for *Chevron* deference in any event. *See* Church of Scientology v. IRS, 792 F.2d 153, 162 n.4 (D.C. Cir. 1986) (en banc) (Scalia, J.), *aff'd*, 484 U.S. 9 (1987). The question of *Chevron* deference is closer in Exemption 3 cases, in which the presence of other statutes can trigger the FOIA exemption. *See, e.g.*, Landmark Legal Found. v. IRS, 267 F.3d 1132, 1134-35 (D.C. Cir. 2001) (holding that I.R.C. § 6103 triggers FOIA's Exemption 3 for tax return information); A. Michael's Piano, Inc. v. FTC, 18 F.3d 138, 143-46 (2d Cir. 1994) (holding that FTC Act § 21(f) triggers FOIA Exemption 3 for information "provided voluntarily in lieu of compulsory process" by parties under investigation); McGilvra v. NTSB, 840 F. Supp. 100, 101-02 (D. Colo. 1993) (holding that 49 U.S.C. § 1905 triggers FOIA Exemption 3 such that the NTSB can not release an unedited version of a cockpit voice recorder tape); *see also* United States v. Mead Corp., 533 U.S. 218, 221 (2001) (limiting *Chevron* deference over an informal agency policy statement and invoking *Skidmore* deference instead).

The de novo standard is the most aggressive judicial review standard available. As hypothesized here, it should yield a higher reversal rate than the more deferential substantial evidence and clearly erroneous standards. Conceptually, even a 50% de novo reversal rate, much like that which occurs before ALJs in SSA disability cases, would not be surprising. But this is clearly not the situation. Over the thirty-five year life of FOIA, it has been estimated that the district court reversal rate is closer to 10% than 50%. In order to confirm the 10% reversal rate estimate, this Article surveyed all FOIA cases decided over the past decade. This study revealed that, of the more than 3600 FOIA cases were decided in the district courts during the ten year period from 1990 to 1999, just over 10% were reversed.[152]

In maintaining this modest reversal rate over such a long time and for so many cases, one has to ask whether the courts have ignored the de novo standard. District courts seem to affirm FOIA cases almost instinctively, and by so doing have produced a real world reversal rate that is closer to the hypothesized arbitrary and capricious standard. In

152. *See* Appendix A. Each FOIA case is given a number in an annual report published by the Department of Justice, Office of Information and Privacy. *See, e.g.*, DOJ FREEDOM OF INFORMATION CASE LIST (2000). Cases were examined as follows: all dispositions that were affirmed, plus voluntary dismissals amounted to 2961 cases. To this 2961 total was added 50% of the 208 partially affirmed and reversed cases on the grounds that they could be called either way; all stipulated dismissals where no attorneys' fees were awarded (546) were also added in. This totaled 3611 cases. Subtracting the cases reversed (189) plus one-half of the cases affirmed/reversed in part (104) produced 293 cases, which when divided by 3611, creating a reversal rate of 8.1%. If the 139 stipulated dismissals are considered to be "cases reversed" when attorneys' fees were awarded (on the assumption that there was merit to the claim), the reversal rate rises to about 11.7%. Thus, for the decade, the reversal rate can be set between 8 and 12%. This suggests that the 10% overall reversal rate for the 36-year life of the program is a reasonable approximation.

effect, the de novo standard in FOIA cases has become the mirror image of the substantial evidence standard in SSA disability cases. If they were switched, the reversal rates might meet outcome expectations. But as they stand, neither standard appears to fulfill the expectations Congress set for them.

Because Congress clearly gave the courts a broad mandate to oversee FOIA cases,[153] one wonders why the invitation has been declined. Again, inarticulate factors seem to offer the only explanations. But first the judicial response must be analyzed to see if it varies depending upon which exemptions are at issue. Although the eight FOIA exemptions are all formally subject to the de novo review standard, some exemptions may engender stricter review than others. That possibility was tested in order to see if it produced a potentially distorting effect on reversal rates.[154]

1. The Special Case of Exemption 1

153. FOIA grants broad jurisdiction to U.S. district courts and provides for in camera examination. 5 U.S.C. § 552(a)(4)(B) (2000). Under FOIA, the agency has the burden of justifying nondisclosure, which it sustains by submitting detailed affidavits that index the documents requested and assert justifications for nonproduction for each document under the claimed exemptions. Affidavits, not denial letters, must be used. *See* DOJ, FOIA GUIDE AND PRIVACY ACT OVERVIEW 592-93 (2000).

154. *See* DOJ v. Julian, 486 U.S. 1, 15 (1988) (Scalia, J., dissenting).

The prime candidate for special treatment is FOIA's Exemption 1, which exempts from disclosure national security information "properly classified pursuant to an Executive order."[155] District courts might be expected to err on the side of nondisclosure of national security information, even when the classifications are questionable. Thus, despite an explicit congressional invitation to do so,[156] the courts have rarely granted requests under this exemption. As a result, even though the de novo standard applies, the judicial instinct remains highly deferential.[157] Prior to 1986, appellate courts had not upheld any decisions to reject an agency's classification claim.[158] A separate analysis of Exemption 1 cases in the 1990s[159] confirms that situation (and in our post-9/11 world we can expect that not to change).

In practice, the de facto standard of review is not "de novo" or even "arbitrary and capricious" in Exemption 1 cases; it is closer to "committed to agency discretion."[160] This review reality is rarely acknowledged; but occasionally a court will admit that Exemption 1 cases are different.[161] The courts seem to have effectively amended the FOIA de novo standard without Congress' concurrence.

But the Exemption 1 phenomenon still does not explain the overall FOIA outcomes divergence. Recalculating FOIA cases without Exemption 1 cases only increases reversals by about 1% to around 11%.[162] Although this reversal rate moves closer to a hypothesized

155. 5 U.S.C. § 552(b)(1); *see* Exec. Order No. 12,958, 3 C.F.R. 333 (1995) (prescribing "a uniform system classifying, safeguarding, and declassifying national security information").

156. In 1974, Congress responded to the case of *EPA v. Mink*, 410 U.S. 73 (1973), in which the Supreme Court held secrets properly classified pursuant to this exemption per se exempt from disclosure, by providing expressly for de novo review and in camera review of documents. *See* 5 U.S.C. § 552(a)(4)(B).

157. *See* Armstrong v. Executive Office of the President, 97 F.3d 575, 580 (D.C. Cir. 1996) (noting the benefits of district courts' "wide discretion" in reviewing FOIA requests).

158. *See* DOJ, FOIA GUIDE & PRIVACY OVERVIEW 83-86 (2000).

159. Cases decided under Exemption 1 during the 1990s were individually reviewed. *See infra* Appendix B. No ultimately successful challenges were revealed. *See, e.g.*, Rosenfeld v. DOJ, 57 F.3d 803 (9th Cir. 1995); Abbotts v. Nuclear Regulatory Comm'n, 766 F.2d 604, 605 (D.C. Cir. 1985) (reversing a district court's finding that requested information did not meet Exemption 1).

160. *See* 5 U.S.C. § 706(2)(A) (2000).

161. In *Stein v. DOJ*, 662 F.2d 1245 (7th Cir. 1981), the court referred to the congressional history of the 1974 Amendments to Exemption 1 and concluded: "Congress did not intend that the courts would make a *true* de novo review of classified documents, that is, a fresh determination of the legitimacy of the classification status of each classified document." *Id.* at 1253 (emphasis added).

162. If Exemption 1 cases are excluded from the denominator, the adjusted reversal rate is only

arbitrary and capricious or substantial evidence reversal rate, it does not approach a "true" de novo rate.[163]

2. Judicial Views About Freedom of Information Act

increased by a percentage point or two. Recalculating 1990-1999 FOIA cases without Exemption 1 cases produces a small increase in reversals to about 11%.

163. *See Stein*, 662 F.2d at 1253.

The Supreme Court was initially supportive of FOIA's purposes,[164] but it has come to view the Act with skepticism, if not resistance.[165] Today, FOIA faces an uphill fight. The Court's approach to FOIA cases began to shift in the 1980s. In *DOJ v. Reporters Comm. for Freedom of the Press*,[166] the Court adopted a more restrictive reading of FOIA in denying access to certain law enforcement records under FOIA Exemption 7(c).[167] The Court determined that FOIA's "central purpose is to ensure that the *Government's* activities be opened to the sharp eye of public scrutiny, not that information about *private citizens* that happens to be in the warehouse of the Government be so disclosed."[168] This judicial gloss[169] encouraged categorical decisions which narrowed the Act's scope.[170]

The Court's skepticism about FOIA is also fed by factors such as the unsympathetic nature of the typical FOIA plaintiff, who need only be "any person,"[171] and the runaway costs of agency compliance.[172] Of course, we can now add another factor to the calculus for determining whether courts should force agencies to produce records: the impact of

---

164. *See, e.g.*, Dep't of the Air Force v. Rose, 425 U.S. 352, 360-61 (1976) ("[FOIA's] basic purpose reflected 'a general philosophy of full agency disclosure unless information is exempted under clearly delineated statutory language.'").

165. As of 1995, twenty-four of the twenty-nine Supreme Court cases ruling on the FOIA have been negative from the requester's perspective. Since 1995, the Court has decided two cases. *See* Dep't of Interior v. Klamath Water Users Protective Ass'n, 532 U.S. 1 (2001); Bibles v. Or. Natural Desert Ass'n., 519 U.S. 355 (1997).

166. 489 U.S. 749 (1989).

167. *Id.* at 774-80.

168. *Id.* at 774.

169. *Cf.* ALEXANDER MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT (1948) (arguing that the purpose of the First Amendment is to protect political speech, not private and commercial speech).

170. *See* Dep't of State v. Ray, 502 U.S. 164, 178-79 (1991) (rejecting disclosure of files of Haitians denied asylum).

171. Unlike sympathetic Social Security disability claimants, FOIA plaintiffs are often prisoners who appear pro se or business competitors seeking to take advantage of the Act for selfish reasons. *See* Sean E. Andrussier, Note, *The Freedom of Information Act in 1990: More Freedom for the Government; Less Information for the Public*, 1991 DUKE L.J. 753, 755-58 (discussing the public interest/personal privacy analysis courts undertake in FOIA cases).

172. *See* DOJ FOIA ANN. REP. (2000), *available at* http://www.usdoj.gov/oip/annual_report/2000/00foiapg9.htm (estimating that FOIA's total cost to the federal government was over $69 million for 2000); DOJ FOIA ANN. REP. (1999), *available at* http://www.usdoj.gov/oip/annual_report/1999/99foiapg25.htm (estimating that FOIA's total cost to the federal government was over $59 million for 1999); *see also* Antonin Scalia, *The Freedom of Information Act Has No Clothes*, REGULATION, Mar.-Apr. 1982, at 14 (discussing FOIA's unintended consequences).

9/11 and concerns about terrorism. The Attorney General has issued a FOIA compliance standard that raises these concerns and counsels against production where there are any doubts.[173]

3. Comparison to Reverse-Freedom of Information Act Cases

A further indicator of attitudes towards FOIA can be gleaned from the judicial reception of "reverse-FOIA" cases, which have been permitted since 1979.[174] In these cases, private parties seek to prevent agencies from *voluntarily* producing documents requested under FOIA. These actions second guess the agency's failure to assert exemptions that might have suppressed production. The judicial review standard for challenging agency action in reverse cases is not de novo, but instead follows the review standards of § 706 of the APA, which usually means arbitrary and capricious.[175]

An outcomes analysis of reverse FOIA cases yields an intriguing statistic: a reversal rate that hovers around 20%.[176] This rate is notable on two counts. First, it is about twice as high as the FOIA de novo rate of 10%. Second, it is close to the range for arbitrary and capricious cases hypothesized here.[177]

Congress considered amending FOIA in 1986 to include coverage of reverse cases.[178] In the course of that effort, the House debated changing the scope of review standard over reverse-FOIA cases to de

173. Memorandum from John Ashcroft, Attorney General, to Heads of All Federal Departments and Agencies (Oct. 12, 2001), *at* http://www.usdoj.gov/04foia/011012.htm (emphasizing the national security and law enforcement interests often at stake in making FOIA decisions).

174. *See* Chrysler Corp. v. Brown, 441 U.S. 281, 285-86 (1979) (authorizing suits to stop agencies from acquiescing in disclosure).

175. Because reverse-FOIA cases are not heard in an on-the-record agency proceeding which would trigger substantial evidence review under § 706(2)(E), the default provision becomes arbitrary and capricious under § 706(2)(A). *See* Citizens to Preserve Overton Park Inc. v. Volpe, 401 U.S. 402 (1971).

176. Reverse-FOIA cases are far fewer in number than FOIA cases. For this Article all reverse cases decided since 1979 were analyzed. Sixty-four reported cases were found. Of that total, forty went to judgment. Thirty-two of those cases were decided in favor of the agency, leaving eight (or 20%) favoring the private objector.

177. *See supra* Part I.C.

178. *See* Freedom of Information Act Amendments of 1986, H.R. 4862, 99th Cong. (1986). House Report 4862 would have given submitters advance notice of agency decisions to release self-designated proprietary information. *Id*. § 2.

novo,[179] which was one of the reasons the bill failed.[180] Ironically, this might have been an unnecessary exercise: based on the statistics analyzed here, submitters are doing better under arbitrary and capricious review than FOIA plaintiffs are doing under de novo review.

179. *See* 132 CONG. REC. 25,172 (1986) (statement of Rep. Weiss).

180. Congressman Weiss stated: "There is no reason that ... [the submitters] need de novo review." *Id.* at 25,173. The point made here is that Congressman Weiss was right in more ways than one.

FOIA review remains a perplexing exercise.[181] The de novo standard of review, which everyone accepts as robust, is revealed as anemic. Even reverse-FOIA cases under arbitrary and capricious review do better in outcomes analysis than de novo FOIA review cases. The black box of "inarticulate factors" seems once again to trump outcomes analysis.

D. Summary

The three different scope of review provisions analyzed above are attached to structurally distinct administrative systems. Social Security disability involves district court substantial evidence review of administrative decision making by ALJs.[182] This is the classic "agency-court" review structure described in *Zurko*.[183] In Veterans

181. A possible explanation might come from the Priest and Klein selection hypothesis, which acknowledges that its 50% win/lose rate does not hold where the defendant has a strong interest in protecting against successful claims. *See supra* notes 31-32 and accompanying text. The DOJ, as the defendant, can effect the reversal rate by not defending weak agency denials and by pulling cases, or, as a last resort, producing the documents sought if the case appears to be a loser before the district judge.

182. The SSA disability system is not formal in the technical sense that it deserves firm application of §§ 554, 556, and 557 of the APA, but it is functionally formal in that an ALJ presides over the hearing. *See* Richardson v. Perales, 402 U.S. 389, 400 (1971) (noting that the conduct of the hearings is generally informal); *see also* Paul R. Verkuil, *A Study of Informal Adjudication Procedures*, 43 U. CHI. L. REV. 739, 739 n.1 (1976) (defining "informal adjudication"). The SSA disability program is one of the few examples of ALJs presiding over hearings not controlled by formal APA procedures.

183. Dickinson v. Zurko, 527 U.S. 150, 152-68 (1999) (discussing at length agency-court and court-court review situations).

Administration disability cases, the clearly erroneous standard is applied by an Article I Court to informal determinations of the BVA, creating an agency-agency review structure. In FOIA cases, the de novo standard is applied by district courts to an informal administrative structure that does not provide a hearing at all. This also results in an agency-court review structure.

The following summary chart ties the actual reversal rates under these three structures to the rates hypothesized earlier:[184]

Chart 2. Actual and Hypothesized Affirmance Rates in SSA, VA, and FOIA Review

Proceeding	Review Standard	Hypothesized Affirmance Rate	Actual Affirmance Rate
SSA-ALJ Review	De Novo	40-50%	50%
SSA-District Court Review	Substantial Evidence	75-85%	50%
VA-CVA	Clearly Erroneous	70-80%	80-85%
FOIA-District Court Review	De Novo	40-50%	90%
Reverse-FOIA-District Court Review	Arbitrary or Capricious	85-90%	80%

This comparison reveals instances where the hypothesis mirrors reality (SSA-ALJ decisions, CVA appeals, and Reverse-FOIA cases), but also instances where it fails dramatically to do so. These are district court decisions in SSA disability and FOIA cases.

184. *See supra* Chart 1, at note 36.

To help explain these anomalies through certain inarticulate factors, the mission, administrative hearing procedures, and scope of review reversal rates might be viewed on a low (L), medium (M), and high (H) scale.[185] The result would look something like this:

	Agency Mission	Formality of Hearing Below	Scope Provision	Review Reversal Rate
SSA	M	H	M	H
VA	H	M	M	M
FOIA	L	L	H	L

The VA disability structure seems the most predictable on this scale, with SSA disability and FOIA decisions still falling outside any notion of predictable outcomes. In sum, of these three examples, only one supports the hypothesis that outcomes actually define standards of review. The next Part will broaden the analysis to include review of sentencing decisions, in order to see whether outcomes analysis might apply to another review standard.

IV. JUDICIAL REVIEW OF SENTENCING DECISIONS

185. This chart uses "agency mission" to define the protective nature of the statute towards its beneficiaries. It refers to the scope of review scale hypothesized in this Article. *See supra* Chart 1, at note 36.

Review of sentencing decisions involves a court-court oversight structure where courts of appeals apply the clearly erroneous standard to district court sentencing decisions. This is court-court review with a twist because it is based on a set of guidelines created by an administrative agency, albeit an agency in the judicial, not executive, branch.[186] Incorporating this structure into the analysis allows us to compare scope of review outcomes in a pure Article III setting, and to inquire whether structuring discretion in advance might be employed at the agency level in order to align outcomes with standards.

Oversight of sentencing decisions is a relatively new venture, one which, like review of veterans disability decisions, is set against a background of unreviewable discretion.[187] Indeed, this shift in accountability is even more dramatic than in the VA review situation because it involves review of Article III institutions for whom independence is the cardinal virtue. Perhaps for this reason many district judges view sentencing oversight as a source of frustration and a diminishment of their role.[188]

---

186. *See* U.S. SENTENCING COMM'N, AN OVERVIEW OF THE U.S. SENTENCING COMM'N (1999), *available at* http://www.ussc.gov/general/ovrvuweb.pdf; *see also* Mistretta v. United States, 488 U.S. 361, 368-70 (1989) (describing the U.S. Sentencing Commission).

187. Until the advent of federal sentencing guidelines, the doctrine of nonreviewability applied to sentencing in the federal court system. *See* Dorszyinski v. United States, 418 U.S. 424, 431-32 (1974). This nonreviewability standard was articulated in *Blockburger v. United States*, 284 U.S. 299, 305 (1932) (stating that only Congress, not the judiciary, may alter a statutory penalty); *see also* Kevin R. Reitz, *Sentencing Guideline Systems and Sentence Appeals: A Comparison of Federal and State Experiences*, 91 NW. U. L. REV. 1441 (1997).

188. *See, e.g.*, Michael Edmund O'Neill, *Abraham's Legacy: An Empirical Assessment of (Nearly) First-Time Offenders in the Federal System*, 42 B.C. L. REV. 291, 339-41 (2001) (discussing the results of a 1996 survey of district court judges).

The Sentencing Guidelines are a complicated structure and they operate under a strict set of rules. The Sentencing Commission uses a "heartland" concept to identify the typical or core case.[189] Inside that core, little sentencing discretion is permitted.[190] Outside the core, district courts may depart from the guidelines, but even in those limited circumstances, judicial review is still active.[191] In *United States v. Koon*,[192] the Supreme Court debated the proper review standard to be applied to departures from the Guidelines. The question was whether the de novo standard or the seemingly less stringent abuse of discretion standard should be applied to Guidelines departures reviewed by the courts of appeals.[193] The Court opted for the latter standard.[194] But even under this standard, which equates to arbitrary and capricious

189. The *Guidelines Manual* directs courts to view each guideline as establishing a "heartland" of typical cases. Departures from the guidelines are meant to be limited to atypical cases. The departures are themselves controlled by a list of permissible and impermissible factors. U.S. SENTENCING COMM'N, GUIDELINES MANUAL, ch. 1, pt. A., Introductory Cmt. 4(b) (2001) [hereinafter GUIDELINES MANUAL].

190. *See id.* ch. 5, pt. A (displaying Sentencing Table).

191. At least one court of appeals has outlined a framework of questions to consider for district courts that want to depart from the Sentencing Guidelines. United States v. Rivera, 994 F.2d 942, 949 (1st Cir. 1993) (Breyer, C.J.). The district court's decision to depart from the guidelines is subject to further review by the court of appeals. *Id.* at 950-52 (describing the review process).

192. 518 U.S. 81 (1996).

193. *Id.* at 91.

194. The *Koon* Court recognized that on questions of law, for example, whether a decision was within the heartland, no deference was owed the district court. *Id.* at 99-100. The Court then found that the district court abused its discretion by considering the defendant's occupation of police officer to justify taking its decision out of the heartland criteria. *Id.* at 110.

review of agency action, the reviewing courts have been alert to unjustified departures from the Guidelines by the district courts.

The reversal rate of appealed cases[195] averages around 20%.[196] This rate is consistent with the arbitrary and capricious reversal rate hypothesized in this Article.[197] This rate of reversals may become constant over time or may simply reflect the transition to a new regime by trial courts learning to adjust. In choosing the abuse of discretion standard over de novo review, the *Koon* Court presumably meant to reinforce the primary role of the district courts in the sentencing process and may have been signaling courts of appeals that higher rates of affirmance should be expected in the future. For purposes here, the crucial point is that the Court chose to send signals to the courts of appeals through the selection of a scope of review standard. Once again, the Court is using scope of review to effect outcomes. Like its decisions on review standards discussed earlier,[198] the Court expects that over time the lower courts or agencies involved will move their outcomes or reversal percentages in the desired direction.

---

195. Only a small percentage of Guidelines cases are appealed. In *Koon*, the Court noted: "In 1994, for example, 93.9% of Guidelines cases were not appealed." *Id*. at 98.

196. U.S. SENTENCING COMM'N, 1999 ANN. REP. 40 (1999) (reflecting 80% affirmance rate), *available at* http://www.ussc.gov/ ANNRPT/1999/ar99toc.htm (last visited Nov. 19, 2002).

197. The universe of appeals is small relative to the total number of cases. That fact undoubtedly affects the reversal rates. It also suggests a large number of cases that fall within the Guidelines are essentially unappealable. By comparison, of almost two million initial SSA claims, in fiscal year 2000, only about 12% reached the Appeals Council (122,780) and only 10% of that total (12,011) reached the federal courts. *See* SSAB, CHARTING THE FUTURE, *supra* note 100 at 8, 21.

198. *See supra* Part II.

The sentencing review experience offers an entirely new ap-proach to review. It establishes a system where initial decisions are made more predictable by a carefully calibrated and elaborate matrix of relevant considerations.[199] This process has lessons for administrative review. By subjecting district courts to oversight of their most fundamental decisions, incarceration of criminals, the Supreme Court has invited an inference: that administrative decisions, those of lesser formality and social importance, could also be susceptible to such control, whether by agencies, Congress, or the courts themselves.

199. The relevant considerations include, but are not limited to, the nature of the offense, the identity of the victim, the defendant's role in the crime, and his or her criminal history. GUIDELINES MANUAL, *supra* note 189, at §§ 2AX, 3A1.1-1.2, 3B1.1-1.4, ch. 4A.

The purpose of the Sentencing Guidelines, to limit discretion by individual judges and thereby enhance uniformity in sentencing,[200] connects to a longstanding problem of administrative law: the use of rulemaking to foreclose factual determinations in adjudication. In the Social Security disability field, the Supreme Court in *Heckler v. Campbell*[201] gave its approval to rules that set medical-vocational guidelines, even though these guidelines had the effect of denying individual ALJ hearings on the issues at stake.[202] The *Heckler* case recognized the enormous potential that rulemaking has for re-calibrating the relationship between the courts and agencies on judicial review. The purpose of the "grid rule" at stake in *Heckler* was to guide agency deciders and to reduce the discretionary component of individualized decision making.[203]

Agencies, which sometimes struggle with congressional and judicial skepticism over their use of generic rules,[204] might learn from the explicit support the Sentencing Commission gives to the rulemaking experience. Structural limitations on sentencing discretion affect fundamental rights by determining the length of prison sentences. By contrast, agency rules created to reduce discretionary decision making will have far less dramatic conse-quences. Moreover, given their experience with the Sentencing Guidelines, district courts—as the reviewers rather than the reviewed in the SSA disability area—might usefully reflect on these contrasting experiences. Reducing discretion through generic rules in light of the sentencing guidelines as an endorsement of earlier agency efforts to regularize disparate decisions by ALJs.[205]

---

200. *See id.* at ch. 1, pt. A.3.

201. 461 U.S. 458 (1983).

202. *Id.* at 468 ("To require the Secretary to relitigate the existence of jobs in the national economy at each hearing would hinder needlessly an already overburdened agency.").

203. *See id.* at 461-62.

204. The distinction between issues that are generic and susceptible to rulemaking and those that are individual and subject to adjudication is an elusive one. This classic dichotomy determines the application of the Due Process Clause. *See* RICHARD J. PIERCE, JR. ET AL., ADMINISTRATIVE LAW AND PROCESS § 6.3.2, at 249-51 (3d ed. 1999) (discussing the individual-group distinction made famous by the *Londoner* and *BiMettalic* cases); *see also* Anaconda Co. v. Ruckelshaus, 482 F.2d 1301, 1306-07 (10th Cir. 1973) (applying the distinction to an EPA rule limiting emissions at one plant).

205. Recently, the SSA has not actively engaged in rulemaking to reduce ALJ discretion in disability adjudications. *See supra* note 21 and accompanying text. The SSA could consider revisiting rulemaking to establish decision parameters by exercising the authority derived from

When it comes to district court review of FOIA decisions, the sentencing experience appears to offer contrary lessons. FOIA decisions are not made by an agency hearing process, let alone one with elaborate rules and procedures regarding the exercise of discretion.[206] Yet, under the far more stringent standard of de novo review, FOIA decisions are routinely affirmed by district courts.[207] The degree of oversight lavished by district courts on Social Security disability cases is absent, as is the care given to reduce judicial discretion in the sentencing context.

It is hard to overcome the impression that in one situation (FOIA) the district courts have failed to grasp the nettle, whereas in the other (SSA disability) they have been reluctant to ungrasp it. The sentencing review experience offers district courts better insights into their reviewing function. This new awareness could be an opportunity to open a dialogue on the scope of review problem with the district courts. The FOIA situation is different. It does not have a specific agency to press its case, nor a court system to favor its potential beneficiaries. All it has is a favorable review standard which is not enough to predict or control decision outcomes.

V. LESSONS IN SCOPE OF REVIEW FOR THE SUPREME COURT AND CONGRESS

It is asking a lot to have scope of review standards reflect outcomes or reversal rates in a predictable way. Review standards have to be measured after the fact, and they are entangled with the inarticulate premises of judicial oversight. Cases have individual characteristics and an unknowable mix of law and facts, such that outcomes are hard to determine in advance. As with umpires, questions of judgment are complicated and calls are rarely obvious.

Heckler v. Campbell. *See supra* notes 201-03 and accompanying text.

206. *See* 5 U.S.C. § 552(6) (2000) (describing how an agency must respond to a FOIA request).

207. *See supra* note 152 and accompanying text.

Despite these difficulties, inquiring about outcomes can be a revealing exercise. The analysis can discern trends and highlight counterintuitive outcomes. Ultimately, the efficacy of a review system is judged by the results it produces. In a broad sense, affirmance, remand, and reversal rates are the results produced. The SSA, VA, and FOIA programs offer opportunities for closer study and action. The question now is: What lessons can be drawn from these divergent scope of review experiences?[208]

A. Reassessing Social Security Administration Disability Review

208. The VA system will not be discussed separately. It is a more recent review structure that, despite criticisms, appears to be implementing the goals Congress set for it. Moreover, it passes the outcomes analysis "test" (20% for clearly erroneous review), so in some rough sense it is in compliance with the goals set here. The VA system remains important, however, if Congress asks whether review of SSA disability cases should be restructured under an Article I system or, conceivably, whether VA disability should be governed by district court review. *See supra* notes 137-38 and accompanying text.

If substantial evidence review of SSA disability cases seems to be more intense than the scope of review provision demands, or even permits, that speaks not so much to subject matter (e.g., whether the claimants are sympathetic[209]) as it does to the agency decision process, including ALJ performance. If the quality of the underlying agency decision matters,[210] and ALJs are our best administrative deciders, then the 50%-plus district court reversal/remand rate for SSA cases challenges conventional wisdom. All three branches should be interested in this question.

1. The Supreme Court's Role: Tinkering with the Substantial Evidence Test

209. In comparison, the VA review system benefits both from an ideal claimant base (the disabled veteran) and from a congressional review scheme that provides for a heightened scope of review standard (clearly erroneous) plus a burden of proof standard that grants the veteran a win in the case of an evidentiary tie (the benefit of the doubt standard). Yet even with these plus factors, VA disability reversal/remand rates are still below those of the SSA. *Compare* Koch & Koplow, *supra* note 102 at 226 (detailing SSA reversal rates), *with supra* notes 143-44 and accompanying text (discussing BVA reversal rates).

210. *See* 1941 ADMIN. PROC. REP., *supra* note 1, at 91 (naming "[t]he character of the administrative agency" and "the confidence which the agency has won" as variables that will influence the reviewing courts).

The Supreme Court has long monitored the procedures surrounding the SSA disability process. In *Richardson v. Perales*,[211] the Court approved the use of written medical reports of doctors over hearsay objections;[212] in *Mathews v. Eldridge*,[213] the Court established a due process balancing of interests test to permit informal pretermination procedures in disability cases;[214] and in *Heckler v. Campbell*,[215] it approved of the "grid" system that re-moved some vocational issues from factual consideration by ALJs.[216] These cases and others,[217] demonstrate the Court's interest in monitoring the procedural aspects of the disability process. It may be time for the Court to take another case.

211. 402 U.S. 389 (1971).
212. *Id.* at 402.
213. 424 U.S. 319 (1976).
214. *Id.* at 334-35 (describing the three inquiries that comprise the balancing test).
215. 461 U.S. 458 (1983).
216. *Id.* at 461-62.
217. *See* Sims v. Apfel, 530 U.S. 103, 112 (2000) (dividing five-to-four over whether issue exhaustion is required in disability cases); Bowen v. Yuckert, 482 U.S. 137, 145 (1987) (upholding SSA's use of a minimum threshold of medical disability in denying benefits).

The next significant SSA disability decision by the Court might deal with the substantial evidence standard of review. It has been more than fifty years since the Court defined the standard under the APA in *Universal Camera*. That case, like the one that should come next, dealt with deference to the fact-based decisions of ALJs, or hearing examiners as they were then called. A new substantial evidence case could be based upon an ALJ decision under the SSA disability process. Such a case could have two goals. It could emphasize the differences among scope of review standards outlined here, and it could help restore substantial evidence as a deferential standard.[218] The Court could also emphasize the connection between accuracy and consistency of result in national programs like SSA disability (or sentencing) by assessing agency attempts to structure administrative discretion through management techniques and rulemaking.

Such a case or cases would address the widely disparate federal judicial outcomes on review of SSA decisions,[219] and help relocate primary responsibility for those decisions back in the hands of the agency and its ALJs. As with the Court's decision in *Koon*, these cases may not have immediate results, but could take hold over time. Moreover, by shaping the dimensions of the substantial evidence test in this context, the Court might also help Congress decide whether it is necessary to consider other legislative alter-natives.

---

218. One line of caselaw that might be reviewed is that of the Second Circuit which held in *Schisler v. Heckler*, 787 F.2d 76 (2d Cir. 1986), that the treating physician's opinion is entitled to controlling weight "unless contradicted by substantial evidence." *Id.* at 81; *see also* Stieberger v. Bowen, 801 F.2d 29, 31-32 (2d Cir. 1986) (citing the *Schisler* summary as the rule in the Second Circuit). This position was later modified by the Second Circuit after an SSA rulemaking. *See* Schisler v. Sullivan, 3 F.3d 563, 567-68 (2d Cir. 1993); *see also* Pierce, *supra* note 42, at 1116-17 (recounting the threat that the Second Circuit's "special rule" posed to the NLRB and the "counter attack" the NLRB made in the form of a legislative rule). Since not all circuits apply it, this rule both modifies the substantial evidence standard and undermines the national reach of the program.

219. The percentage of judgments on the merits won by SSA claimants varies greatly by judicial district, ranging from 2.3% in the Eastern District of Kentucky to 59.8% in the Eastern District of New York. *See* Paul Verkuil & Jeffrey Lubbers, *Alternative Approaches to Judicial Review of Social Security Disability Cases: A Report to the Social Security Advisory Board*, App-A (2002), *available at* http://www.ssab.gov (last visited Nov. 19, 2002).

2. The Role of Congress: Changing the Structure of Review

Pressure is building for Congress to consider legislative alter-natives to district court review of disability cases.[220] Both the Long Range Plan for the Federal Courts and the recommendations of the Social Security Advisory Board raise the question of Article I review. The Judicial Conference proposals postulate an Article I review structure with limited judicial review along the lines of the Veterans' Administration disability program,[221] and the Advisory Board discussed both Article I (Social Security Court) and Article III (Court of Appeals for Social Security) alternatives.[222] At some point Congress may feel compelled to act.[223]

220. *See id.* at 57-59 (outlining proposals for an Article I at Social Security Court by Congress and the Department of Justice). For arguments on both sides of the Article I Court proposal, see *id.* at 61-62.

221. *See* JUDICIAL CONFERENCE OF THE UNITED STATES, LONG RANGE PLAN FOR THE FEDERAL COURTS 34 (1995).

222. SSAB, CHARTING THE FUTURE, *supra* note 100, at 61-62. The SSAB ultimately concluded that it would not recommend an Article III Court of Appeals, but would "favor serious consideration of an Article I Social Security Court." *Id.* at 67.

223. The Ways and Means Social Security Subcommittee has recently considered such changes. *See* Social Security Court of Appeals Act of 1995, H.R. 1587, 104th Cong. (1995); Social Security Procedural Improvements Act of 1993, H.R. 3487, 103d Cong. (1993); Social Security Court of Appeals Act, H.R. 3265, 103d Cong. (1993); Social Security Procedural Improvements Act of 1991, H.R. 2159, 102d Cong. (1991); Social Security Procedural Improvements Act of 1989, H.R. 2349, 101st Cong. (1989).

SSA disability cases constitute a significant share of all federal cases. During the twelve-month period ending September 30, 1999, disability insurance and supplemental security income cases constituted about 5.9% of all federal civil cases terminated by court action.[224] SSA disability cases are the largest category of cases against the United States in the district courts, so if the federal courts are looking to reduce their dockets, these are likely candidates. However, the amount of district court trial time devoted to these cases has been reduced due to the trend toward remands rather than reversals[225] and by the fact that district judges utilize their magistrates to decide many SSA disability cases, which further conserves judicial time.[226]

An Article I disability court system is of primary appeal from the perspective of consistency of outcomes. Consistency is the greatest need in mass justice situations where accuracy per se is a far more elusive quest. Federal court review, on the other hand, is balkanized; it can send inconsistent messages on the law and frustrate evenhanded review on the facts. A single reviewing body can set national standards much like the BVA has done for VA disability cases.[227]

There will be resistance to such a change.[228] Claimants, most of whom are now represented by attorneys,[229] could be a strong po-tential

---

224. FEDERAL JUDICIAL CENTER, JUDICIAL BUSINESS OF THE U.S. COURTS: 1999 ANNUAL REPORT OF THE DIRECTOR, tbl. C-4 at 160-62 (1999) [hereinafter FEDERAL JUDICIARY CTR., 1999 ANNUAL REPORT], *available at* http://www.uscourts.gov/judbus1999/contents.html. There were 228,190 civil cases terminated of which 13,451 were Social Security disability cases.

225. In 1999, only about 0.15% of disability cases reached trial. Federal courts have been backing away from all actual trials during this period as well. *See* Resnik, *supra* note 105, at 925 (documenting and lamenting a federal court trial rate of 8% in 1994). This trend has continued. As of September 30, 1999, the federal civil trial rate had reached 2.3%. FEDERAL JUDICIARY CTR., 1999 ANNUAL REPORT, *supra* note 224, at 160.

226. Magistrates now decide over 40% of disability cases. This fact cuts two ways on the question of Article III review, however, because magistrates are in effect Article I deciders whose credentials are like those of ALJs.

227. The Social Security Article I Court idea is in fact premised upon the capacity to produce "greater uniformity." SSAB, CHARTING THE FUTURE, *supra* note 100, at 23.

228. There should be no constitutional problem with Article I review, so long as Article III courts retain jurisdiction over legal issues as the Federal Circuit does with VA disability review. *See supra* note 137. Shifting from Article III review to Article I review, however, is a much bigger step politically than changing from no review to Article I review as in the VA disability situation.

229. SOCIAL SECURITY ADMINISTRATION, APPROVAL OF CLAIMANT REPRESENTATIVES AND FEES PAID TO ATTORNEYS (2001), *available at* http://www.ssa.gov/oig/adobepdf/A-12-00-10027.pdf ("In Fiscal Year 2000, about 75% of disability claims had attorney representation."). The SSAB has estimated that annual attorney fee payments in disability cases is over $500 million. *See* SSAB, CHARTING THE FUTURE, *supra* note 100, at 2. This could create strong resistance to changing the

lobby against Article I review status for SSA disability claims. Moreover, some of the uniformity argument can be addressed through the creation of a special Article III appeals court. This alternative, which is undesirable for other reasons, might make an administrative structure a harder sell. Before trying to reconceptualize judicial review, however, Congress may be interested in whether the Supreme Court can unify district courts through a "rejuvenated" substantial evidence test.

3. Role of the Agency: Managing, Not Dictating, Outcomes

The Social Security Administration Office of Hearings and Appeals (OHA) bears management responsibility for the disability decision process. As we have discussed, during the 1980s, the OHA aggressively sought to control the decisions of ALJs through their own motion review of those judges with high reversal rates and through the setting of caseload requirements.[230] After strong judicial and congressional opposition to its tactics, OHA backed off its active monitoring of ALJ performance. Today it does little to manage the workload or outcomes of ALJs.[231] This has created something of a policy vacuum. Expert policy leadership seems to have been assumed by the Social Security Advisory Board (SSAB). The SSAB was created as an oversight body when the SSA became an independent agency in March of 1995.[232] But the SSAB, though effective at the conceptual level, is not in a position to carry out actual management reforms.

---

system unless comparable payments are available under the new system.

230. *See supra* notes 124-29 and accompanying text (discussing the Bellmon Amendment).

231. *See supra* note 128 and accompanying text.

232. SSA, *Social Security Online: History Page*, *available at* http://www.ssa.gov/history/keydates.html (last visited Nov. 26, 2001).

The SSAB has placed valuable emphasis upon the need for reform of the disability hearing process,[233] especially by improving the SSA-ALJ relationship.[234] It has also sought to rationalize the role and workload of the Appeals Council, and to improve the management of the state field offices that make the initial decisions.[235] Strengthening the federal-state relationship in man-aging the disability decision process is overdue. Hiring practices among states need to be regularized and training and quality assurance procedures improved.[236] Federal guidelines improving policy implementation need to be made controlling over state deciders (and ALJs for that matter). Policy guidance would allow the SSA to improve on the norms of consistency and uniformity and, by so doing, gain the confidence of the reviewing courts.

B. Reassessing Freedom of Information Act Review

FOIA cases are hard if not impossible to explain in terms of outcomes analysis if de novo is to be a meaningful standard of review. An affirmance rate of almost 90% is indicative of a system of review that does not inquire deeply into the underlying administrative action, especially since there is no hearing below.[237]

1. The Supreme Court's Role in Setting the Tone

233. SSAB, CHARTING THE FUTURE, *supra* note 100, at 18-22.

234. The Board proposes to have the agency represented at the ALJ hearing, to close the record after the ALJ hearing, and to certify claimant representatives according to a system to be established by federal rules. *Id*. at 19-21. These are promising alternatives whose economic effects need to be calculated.

235. *See id*. at 16-17 (outlining proposals to improve communication between the SSA and state agencies and to make all states conform to federal standards).

236. *Id*. at 17. The federal government pays 100% of the costs of the program and could take over the process. *Id*. at 16. That suggestion may be impractical in the short run given the number of deciders involved and may be politically unattractive in the long run for the same reasons. In the interest of avoiding delays in the process, one suggestion being actively considered by the SSA is the elimination of the state reconsideration stage. *Id*. at 17-18. Of course, given the recent fight in Congress over federalizing the airline security work force, it would not be easy to convert state actors in disability cases into federal employees. *See* Juliet Eilperin & Ellen Nakashima, *Airport Security Accord Reached; Measure Calls for Federal Screening Force*, WASH. POST, Nov. 16, 2001, at A1.

237. *See supra* note 152 and accompanying text.

Searching review of FOIA cases depends upon the presence of the de novo standard itself. Unless the Court becomes interested in invigorating that standard, or even in acknowledging it, there is not much point for it to act. Unlike Social Security disability review, where the Court's guidance might be expected to moderate the force of judicial intervention, a Court pronouncement on FOIA review would likely have the opposite effect. The challenge is to make district judges more curious about these cases so that they might look behind agency affidavits. At this juncture, any change in that regard is up to the other branches.

2. Congressional Alternatives

Congress could explore several ways to enhance review of FOIA cases. The first is legislatively to acknowledge that the national security exemption is never going to be subjected to the scrutiny contemplated by de novo review.[238] Amending the review standard to subject national security review to the arbitrary and capricious standard once the classification process is completed would con-form the review standard to reality. Indeed, in the tense security environment we all face, such a modification seems particularly justified.

Moreover, by acknowledging distinctions among the exemptions in terms of review standards, the remaining exemptions might achieve invigorated review simply by comparison. The exemptions that retain de novo review might well be given a closer look.

Congress could also create an administrative review process that would support and even supplant much of the work of the district courts. An agency review process, staffed either by ALJs or other qualified agency deciders, could couple the de novo standard with increased in camera review. As Professor Grunewald has suggested, this agency process could be operated by the Department of Justice either on a consent basis or upon referrals from the district courts.[239]

238. *See supra* notes 160-61 and accompanying text (describing the unique quality of Exemption 1 cases).

239. *See* Mark H. Grunewald, *Freedom of Information Act Dispute Resolution*, 40 ADMIN. L. REV. 1, 1-3 (1988) (describing administrative alternatives to court review in FOIA cases).

These decisions might then be directly reviewed in the courts of appeals,[240] thereby reducing burdens of district courts.

A less ambitious alternative would be to create a conciliation or ombudsman function to be administered by a separate admin-istrative agency created to resolve FOIA cases or by the Department of Justice itself.[241] Because this approach would represent a shift in emphasis away from the courts, Congress may not want to go this far. Indeed, the requester community would likely resist any change that moves away from de novo review. Finally, Congress could use its oversight function to explore with agencies their approach to FOIA review. By selecting the agencies with the higher FOIA request caseload, some efforts could be made to ensure that these agencies give their requests the prompt and fair responses they deserve.

3. Role of Agencies and the Department of Justice in Freedom of Information Act Review

Much of what Congress might achieve at the agency level could be achieved by the agencies themselves with assistance from the Department of Justice (DOJ). One possibility would be to formalize the FOIA ombudsman function either within agencies or within the DOJ. The DOJ Office of Information and Privacy, which is responsible for collecting statistics on FOIA cases, claims to support the ombudsman concept.[242] But there is little evidence that the public is aware of the internal review possibility. The agency receives only six to twelve requests for ombudsman review per year,[243] which is a tiny portion of the nearly two million requests agencies receive annually or even of the

240. *Id*. at 43-44.

241. *See id*. at 45-48. Professor Grunwald concludes that "[t]he ombudsman function proposed alternatively offers a meaningful prospect for meeting [the] needs" of review "in many cases short of resort to adjudication but without restriction of the existing judicial remedy." *Id*. at 65.

242. The DOJ website cites the study by Professor Grunewald on which his article was based. DOJ, *Focus on FOIA "Ombudsman" Role*, DOJ FOIA UPDATE, Fall 1987, *available at* http://www.usdoj.gov/oip/foia_updates/Vol_VIII_3/viii3page2.htm (approving of the ombudsman role, but noting that no formal mechanism has been adopted).

243. Discussion with Dan Metcalfe, Co-Director, DOJ Office of Information and Privacy (Oct. 20, 2001).

approximately one million that are denied.[244] A diligent search of the DOJ website fails to reveal this function and groups that monitor FOIA decisions for claimants, like Public Citizen, are unaware that the service exists.[245] The DOJ Office of Information and Privacy could easily formalize the ombudsman role within the agency and make it more generally available. Of course, the ombudsman cannot correct ongoing cases, but it can discern trends and problem areas at particular agencies and recommend corrections to the process that can serve to avoid the necessity for district court review.[246]

CONCLUSION

This Article introduces a concept of outcomes analysis as a way to understand and critique scope of review standards and to determine their impact upon administrative and judicial behavior. By postulating affirmance/reversal formulas for deciding cases under the various standards, it presents a purely hypothetical construct. Still, studying outcomes in selected high volume settings probes whether Congress' will is being followed and helps to establish predictive bases for deciding cases.

There is much that can be learned from this highly academic exercise. The Court, Congress, and the agencies can use these results to stimulate discussion and even to reconcile the application of these standards to specific review situations. All of this can be done while stopping far short of some mechanical application. Judges are not automata, and cases are not sausages, but consistency of application is still a worthy goal. This makes the question of outcomes worth pursuing even—especially—some fifty years after scope of review standards were first given new articulation under the APA.

---

244. *See supra* note 96 and accompanying text.

245. Discussion with David Vladeck, Director of Litigation Group, Public Citizen (Oct. 25, 2001).

246. Cases that get dismissed before the district courts often result in the production of requested documents. *See supra* note 152 and accompanying text. Under an ombudsman regime, these cases need not be brought in the first place.

APPENDIX A

Year	Disposition Affirmed	Disposition Reversed	Voluntary Dismissal	Stipulated Dismissal	Aff. in Part Rev. in Part
1990	251	17	32	62	26
1991	265	24	8	63	33
1992	204	15	13	42	13
1993	273	23	14	75	38
1994	286	17	20	98	18
1995	273	26	32	72	13
1996	275	14	27	61	18
1997	168	16	20	65	18
1998	230	20	13	74	15
1999	192	17	20	73	16
Total	**2417**	**189**	**199**	**685**[247]	**208**

Year	Dismissed as Moot	Dismissed Sua Sponte	Failure to Prosecute	Failure to Serve	Dismissed by Court
1990	9	4	8	3	1
1991	9	2	16	3	2
1992	10	1	-	1	1
1993	9	5	7	1	1
1994	31	4	3	2	4
1995	9	7	13	-	2
1996	14	3	10	-	4
1997	8	3	2	-	1
1998	8	3	2	-	-
1999	5	2	6	-	1
Total	**112**	**34**	**67**	**10**	**17**

Year	**Exhaust Admin. Rem.**	**Failure to Name Party**	**Jurisdictional Defect**	**Dismissed Res Judicata**	**Failure to State a Claim**
1990	**12**	**-**	**1**	**-**	**-**
1991	**10**	**1**	**2**	**-**	**-**
1992	**4**	**2**	**2**	**-**	**2**
1993	**7**	**1**	**2**	**1**	**-**
1994	**9**	**1**	**1**	**-**	**2**
1995	**13**	**3**	**-**	**1**	**-**

247. 139 of the 685 stipulated dismissals were with costs.

1996	**7**	**-**	**-**	**1**	**-**
1997	**4**	**-**	**1**	**-**	**-**
1998	**5**	**-**	**3**	**-**	**-**
1999	**7**	**-**	**-**	**-**	**-**
Total	**78**	**8**	**12**	**3**	**4**

Dis. Aff. + Vol. Dis. + Pro. Dis. + Stip. Dis. + 1/2 in part: 3611
Dis. Rev. + 1/2 in part + Stip. Dis. w/ costs: 432
Total Cases: 4043

(432 of 4043 equals a 10.7% reversal rate)

APPENDIX B

Exemption 1 Cases

Ex. 1 was the only exemption:	50 Dispositions Affirmed	(3561 total)
Ex. 1 was one of several exemptions:	126 Dispositions Affirmed 24 Aff in part/Rev. in part 14 Stipulated Dismissals	
Total Ex. 1 cases:	214 Dispositions Affirmed	(3397 total)

Reversal rate excludes all cases where Ex. 1 was the only exemption: 432/3993 = 10.8%

Reversal rate excludes all cases where Ex. 1 was one of several exemptions: 432/3829 = 11.3%