

2016

Changing Standards of Review

Jeffrey C. Dobbins

Follow this and additional works at: <https://lawcommons.luc.edu/lucj>



Part of the [Law Commons](#)

Recommended Citation

Jeffrey C. Dobbins, *Changing Standards of Review*, 48 Loy. U. Chi. L. J. 205 ().

Available at: <https://lawcommons.luc.edu/lucj/vol48/iss1/6>

This Article is brought to you for free and open access by LAW eCommons. It has been accepted for inclusion in Loyola University Chicago Law Journal by an authorized editor of LAW eCommons. For more information, please contact law-library@luc.edu.

Changing Standards of Review

*Jeffrey C. Dobbins**

Do standards of review matter? On the one hand, judges insist that they do, and appellate practitioners know that they ignore standards of review at their peril. On the other hand, it is not unusual to find judges and academics who concede that there is not much difference between the many standards of review, and that the articulation of the standard may not make much of a difference in reversal rates.

To test the question, researchers would ideally take a set of cases, have a court decide them under one standard of review, and then, while avoiding any bias from deciding the same cases twice, have the same court decide the cases under a different standard. Though this is not a practical experiment, recent legal history offers examples in which legislative and judicial decisions have altered the standard of review for particular types of cases. By comparing reversal rates that precede a change in the standard against reversal rates in the same court from cases that post-date it, researchers can come as close as possible to the ideal experiment: testing the same cases under two different standards.

An initial look at one instance of a changed standard of review supports the proposition that merely altering the standard does not shift reversal rates. This look also suggests a hypothesis about standards of review that might reconcile the tension between those who are convinced that standards matter and those who think they do not. In particular, this Article suggests that standards of review are unlikely to drive changes in reversal rates because the standards are not drivers; rather, standards reflect an understanding about the role of various institutions within the legal system in a particular type of case. Unless those understandings change, reversal rates are unlikely to change even if the standard does. For that reason, legislatures and practitioners seeking to influence appellate court reversal rates should focus more of their attention on changing those understandings, rather than merely changing the language of the standard. Finally, this Article suggests an agenda for future investigations into the standard of review by pointing out two other

* Associate Professor of Law, Willamette University College of Law.

instances where the standard changed. Examining reversal rates associated with those instances will continue this investigation into the source and significance of standards of review in our legal system.

INTRODUCTION.....	207
I. STANDARDS OF REVIEW: FUNDAMENTALS	214
A. <i>Current Understanding</i>	214
1. Appellate Court Review of Trial Court Decisions	215
2. Review of Administrative Decisions	217
3. Appellate Court Control over Standards of Review	219
B. <i>History</i>	224
C. <i>Purposes</i>	226
D. <i>Does a Different Standard Affect Outcomes?</i>	227
1. Different Standards Affect Outcomes (Circumstantial Evidence)	227
2. Different Standards Affect Outcomes (Empirical Findings)	229
3. Different Standards Do Not Affect Outcomes (Circumstantial Evidence)	232
4. Different Standards Do Not Affect Outcomes (Empirical Findings)	234
5. Additional Evidence Regarding the Effect and Nature of Appellate Standards of Review	236
II. CHANGING STANDARDS OF REVIEW: THREE EXAMPLES	237
A. <i>Discretionary De Novo Review in Equitable Actions or Proceedings in Oregon, 2005–2014</i>	237
B. <i>Other Examples of Changing Standards of Review</i>	243
1. Review of Interpretations of Law by Federal Administrative Agencies, 1987–1997, 2002– 2012.....	243
2. Review of Downward Departures Under the Federal Sentencing Guidelines, 1991–1997	247
III. LESSONS FROM CHANGE AND SUGGESTIONS FOR CHANGE	249

INTRODUCTION

The bread and butter of appellate advocacy is the standard of review.¹ From Supreme Court Justices to professional instructors of legal writing,² from bar journals to hornbooks on the appellate process,³ those who offer their advice to appellate counsel emphasize the importance of phrasing one's argument in terms of the applicable standard.

Ask most appellate judges and practitioners, and they will confirm that this emphasis on the standard of review is not mere rhetorical flourish. The standard of review that governs a particular issue on appeal has been characterized as the pivot on which judicial decisions turn.⁴ D.C. Circuit Judge Patricia Wald once wrote that the standard of review in a case "more often than not determines the outcome" of that appeal,⁵ and Tenth Circuit Judge Deanelle Tacha wrote that "to the [appellate] judge, [the

1. The standard of review defines how deferential an appellate court should be to the judge (or agency) under review; as such, it effectively determines how willing an appellate court should be to reverse the decision below. See DANIEL J. MEADOR ET AL., *APPELLATE COURTS: STRUCTURES, FUNCTIONS, PROCESSES, AND PERSONNEL* 222 (2d Ed. 2006) ("[S]tandards of review address ... whether the reviewing court should defer to the trial court and, if so, to what extent."). For instance, under a deferential standard of review, such as "clear error," an appellate court would reverse only if it was strongly convinced that the prior decision was in error, and reversal should be relatively rare. MICHAEL D. MURRAY & CHRISTY H. DESANCTIS, *APPELLATE ADVOCACY AND MOOT COURT* 5–6 (2006). Under less deferential standards of review, such as *de novo* review, an appellate court makes up its own mind, and gives no particular weight to the decision below, so reversal should be more common than under deferential standards. *Id.* ("The difference between an issue on appeal case that is governed by a *de novo* standard . . . as opposed to an issue governed by a clearly erroneous standard . . . is the difference between an appeal that may have a decent change and one that may have a snowball's chance in hell."). See also Paul Verkuil, *An Outcomes Analysis of Scope of Review Standards*, 44 WM. & MARY L. REV. 687–88 (2002); Part I.A.1 (describing the process of appellate review and the three standards that this Article will specifically contemplate).

2. See BRYAN A. GARNER & ANTONIN SCALIA, *MAKING YOUR CASE: THE ART OF PERSUADING JUDGES* 11 (2008) (noting that the fifth "general principle of argumentation" is to "pay careful attention to the standard of decision").

3. See Henry Deeb Gabriel, *Unequaled Expertise: Childress and Davis' Federal Standards of Review*, 12 J. APP. PRAC. & PROCESS 163, 164 ("[T]he appropriate standard of review is essential to appellate practice."); Nathan Hecht, *Revisiting Standards of Review in Civil Appeals*, 24 ST. MARY'S L.J. 1041, 1041 (1993) (noting the standard of review is "the anvil on which the recitation of facts and the argument are both to be forged, and it as much as anything else determines their cast. Each word should be aimed at proving that the ruling on appeal either does or does not meet the standard by which it is to be judged").

4. MICHAEL D. MURRAY & CHRISTY H. DESANCTIS, *ADVERSARIAL LEGAL WRITING AND ORAL ARGUMENT* 103 (2006) ("The standard of review can have a tremendous impact on the chances of success on an appeal.").

5. Patricia Wald, *The Rhetoric of Results and the Results of Rhetoric*, 62 U. CHI. L. REV. 1371, 1391 (1995).

standard of review] is everything.”⁶

The importance of the standard in appellate practice, then, is difficult to overstate. By setting the decision rule for when the appellate court is able to reverse the findings below, the standard effectively calibrates an “index point” for the likelihood of reversal, as determined by a level of appellate court confidence in the validity of the ruling below. The standard, therefore, serves to define the relationship between the reviewing court and the entity whose decision is under review. It outlines the scope of authority that reviewing judges have over the decision, and it often reflects a statutory—and occasionally, a constitutional—mandate.⁷ Those who offer advice to appellate attorneys invariably focus their attention on the standard of review in appellate courts, and on the impact that the standard has on the appellate process.

Despite the importance of the standard of review, however, the source of the standard in a given case is often curiously underdetermined by positive law and “imperfectly provided by precedent.”⁸ It is true that rules and statutes occasionally define the relevant standard.⁹ In other cases, though, no law provides the standard. In an effort to determine what the standard of review should be, the appellate court and counsel may engage in a complicated analysis involving conclusions about the “factual” or “legal” nature of the question presented, as well as the relevant roles of appellate and trial courts with respect to that question.¹⁰

6. HARRY T. EDWARDS & LINDA A. ELLIOTT, *FEDERAL STANDARDS OF REVIEW: APPELLATE COURT REVIEW OF DISTRICT COURT DECISIONS AND AGENCY ACTIONS*, at vii (2005).

7. *See, e.g.*, U.S. CONST. AMEND. VII (“[N]o fact tried by a jury, shall be otherwise re-examined in any court of the United States, than according to the rules of the common law.”); 5 U.S.C. § 706 (2015) (defining scope of review); FED. R. CIV. P. 52(a)(6) (noting judicial findings of fact “must not be set aside unless clearly erroneous”).

8. Steven Alan Childress, *Constitutional Fact and Process: A First Amendment Model of Censorial Discretion*, 70 TUL. L. REV. 1229, 1232–33 (1996); *see also* Cooter & Gell v. Hartmarx Corp., 496 U.S. 384, 401 (1990) (noting the difficulty of first order determination between issues of fact and law).

9. This is particularly true in the review of administrative agency decisions, given the prevalence of statutory Administrative Procedure Acts. *See, e.g.*, 5 U.S.C. § 706 (describing the scope of review for reviewing courts); OR. REV. STAT. § 183.400 (2016) (describing the judicial review of agency rulemaking); OR. REV. STAT. § 180.482 (2016) (reviewing agency orders for contested cases); OR. REV. STAT. § 180.484 (2016) (reviewing agency orders for cases other than contested cases).

10. Consider, for instance, the multi-page discussions regarding the appropriate standard of review in *Sch. Dist. of Wisconsin Dells v. Z.S., ex rel. Littlegeorge*, 295 F.3d 671, 674–75 (7th Cir. 2002) (discussing the appropriate standard of review in a challenge to a state Administrative Law Judge (“ALJ”) decision regarding violation of Individuals with Disabilities Education Act), in *State v. Navas*, 913 P.3d 39, 49 (Haw. 1996) (discussing the standard of review for determinations of probable cause in an issuance of a search warrant), in *Mars Steel Corp. v. Cont’l Bank, N.A.*, 880 F.2d 928, 939 (7th Cir. 1989) (discussing the standard for review for the imposition of Federal Rule

It is rather surprising that such a critical part of the appellate process—determining the intensity with which appellate courts review decision making on a particular topic—remains so often undetermined.¹¹

Furthermore, in the midst of this substantially unguided effort to define the scope of review, there is a curious conflict at the core of appellate standard of review definitions: advocates, judges, and academics repeatedly suggest that the standard might not affect outcomes as much as one might think. It is, for example, not at all unusual for appellate judges to hint—amid discussions on the topic—that the particular standard of review may not really matter.¹² Similarly, advocates calling for a particular standard of review might have difficulty identifying circumstances when a different standard would really make a difference in the outcome of their case.¹³ And often, appellate courts brush aside difficult problems in defining the standard of review by noting that one standard or another would likely not make a difference in a particular case.¹⁴

This case-specific skepticism about whether a particular standard of review actually affects the likelihood of reversal is not clearly resolved by academic studies where researchers have examined the impact of the standard of review on appellate outcomes. On the one hand, some empirical investigations suggest that the standard of review has the predictive value that one might expect—noting that deferential standards lead to reversals less often than non-deferential standards.¹⁵ On the other

of Civil Procedure 11 sanctions), and in *Cooter & Gell*, 496 U.S. at 405 (same).

11. To this extent, the standard of review is like other areas of appellate practice—the rules of precedent, the record review rule and exceptions to it, and the exercise of appellate courts’ inherent and supervisory authority—that amounts to a “procedural common law” for appellate courts. See generally Jeffrey C. Dobbins, *New Evidence on Appeal*, 96 MINN. L. REV. 2016 (2012) (discussing the underlying principles in appellate record review); Jeffrey C. Dobbins, *Structure and Precedent*, 108 MICH. L. REV. 1453 (2010) [hereinafter Dobbins, *Structure*] (highlighting the difficulties in evaluating the precedential effect when the standard structure or process of appellate review is altered).

12. See *United States v. Boyd*, 55 F.3d 239, 242 (7th Cir. 1995) (“[T]here are more verbal formulas for the scope of appellate review . . . than there are distinctions actually capable of being drawn in the practice of appellate review.”).

13. See, e.g., Transcript of Oral Argument at 40, *Metro. Life Ins. Co. v. Glenn*, 554 U.S. 105 (2008) (No. 06-923) (citing to Justice Alito: “I don’t know how thin you can slice these standards of review”).

14. See *Dickinson v. Zurko*, 527 U.S. 150, 163 (1999) (“[W]e have failed to uncover a single instance in which a reviewing court conceded that the use of one standard rather than the other would in fact have produced a different outcome.”).

15. See, e.g., Corey Rayburn Yung, *Judged by the Company You Keep: An Empirical Study of the Ideologies of Judges on the United States Courts of Appeals*, 51 B.C. L. REV. 1133, 1159 (2010) (“[S]tandards of review could increase or decrease the likelihood of reversal of district court

hand, equally thoughtful studies have concluded that the typical spectrum of standards of review imperfectly reflects—and, occasionally, fails to predict—relevant rates of reversal.¹⁶ One would expect, for instance, that reversal rates should be lower when appellate courts use a highly deferential standard of review (like arbitrary and capricious review) than when they use a more skeptical standard (such as *de novo* review). That expectation, however, does not always pan out in practice.¹⁷ So it appears that, at least in part, the skeptics are right: sometimes, the standard of review makes little difference to reversal rates.

How can the apparent importance of the standard of review to the appellate process be reconciled with evidence and anecdotes suggesting that the precise standard of review does not matter to the degree that one might expect?

In an effort to address this question, this Article examines the role of the standard of review in the state and federal courts of the United States. Part I briefly describes the current understanding of traditional standards of review: what they are, the history of their development, and the consensus understanding regarding their role in the American legal system. Part I also incorporates a discussion of the standard of review in administrative law contexts. While there are, of course, differences between appellate court review of trial court decisions and the role of courts in reviewing administrative agencies' decisions, the relevant standards are often discussed together, or in comparison with one another, and the concerns and considerations regarding the importance of standards of review are similar in the two contexts.

Part I then continues by examining the current understanding—both anecdotal and empirical—about the degree to which the standard of review matters (i.e., whether the standard does, in fact, make a difference in the outcome of cases under review). As already noted above, the jury is out (so to speak) on this question: the data suggests that the standard might not make as much of a difference in case outcomes as the language of various standards might suggest.

This seems odd, indeed. Given the importance of standards of review in the work of appellate judges and appellate practitioners, one would

judgments"); see *infra* Part I.D (ascertaining whether different standards affect outcomes).

16. See Paul Verkuil, *An Outcomes Analysis of Scope of Review Standards*, 44 WM. & MARY L. REV. 679, 721 (2002) (“[R]ate of reversals may become constant over time or may simply reflect the transition to a new regime by trial courts learning to adjust.”).

17. See *id.* at 687, 717 (finding that Freedom of Information Act reversal rates under the arbitrary and capricious review standard did not correlate with the hypothesized difference predicted by the author).

hope that a different standard of review would make a predictable difference in the rate of reversal—a critical measure of appellate review. If a difference in the standard does *not* make a predictable difference in the rate of review, then what is the point of the hue and cry regarding the standard of review at all?¹⁸ How can one reconcile the conviction of practitioners and judges—as well as the cost in time and energy—regarding the fundamental importance of the standard of review, with the empirical suggestion that differing standards have only incidental impact on the result of cases in the judicial process?

Existing literature contributes to the conflict as much as it helps to reconcile it. Current empirical investigations are necessarily complicated by the multitude of decision makers and the variety of causes of action under review in the pool of relevant decisions. While the large number of cases reviewed allows investigators to make useful statements about *overall* reversal rates,¹⁹ the investigations are necessarily hampered by the fact that any comparison of reversal rates between different standards of review also involves cases on appeal that differ in type.²⁰ In such instances, any differences in reversal rates might be explained by different judicial attitudes toward the underlying substance of the appeal, rather than the different approaches to the standard of review.

In an effort to avoid this problem and to extend the existing literature, Part II of this Article takes a different approach. Rather than comparing reversal rates under different standards of review as applied at the same moment in time over different types of cases, this Article examines circumstances in which the standard of review changed over time for the *same type* of cases. By limiting the investigation to a particular category of cases in which such a change occurred, this approach limits any variation in reversal rates that might result from the difference in the type of case, rather than the standard of review.

This approach is still imperfect, of course. It trades variation caused by differences in case types for variation associated with changes over

18. In the absence of consensus regarding the value (or lack thereof) of careful attention to the standard of review, judges and counsel might well still write opinions and briefs targeted to standards of review; if the standards were only rhetorical window-dressing, though, we might expect parties to ultimately fall out of practice.

19. This is essentially the conclusion in these studies. See, e.g., Yung, *supra* note 15, at 1159, 1161–65 (discussing how adjusting for standard of review in an empirical study can determine how appellate judges' ideology influences reversal rates).

20. For instance, Verkuil's examination into the difference between reversal rates under an "arbitrary and capricious" standard of review and reversal rates under a "de novo" standard required comparing administrative appeals under the Social Security Act, on one hand, and under FOIA, on the other. Verkuil, *supra* note 16, at 689, 710–14. Such fundamental differences in case type can easily complicate the analysis about the importance of standards of review. *Id.*

time. As time passes, the judges on the relevant court will change, as do facts between cases—even if the cases are of the same type. But this complicating factor is not absent from existing studies: there was no guarantee that the same panels of judges were deciding cases or that the facts were constant between cases with different standards. Under this Article’s approach, the pool of decisions under review after a change in the standard of review should be statistically similar in makeup to the pool of decisions that preceded the relevant change. This similarity should allow a relatively direct comparison between reversal rates in the pre-change and post-change eras.

Part II focuses its attention on one instance in recent legal history in which the standard of review changed for a particular class of cases. Prior to 2009, the Oregon Court of Appeals was required to review facts found in “equitable proceedings” under a *de novo* standard. In 2009, however, the Oregon Legislative Assembly amended the relevant statute and gave the Court of Appeals the discretion to use either *de novo* review or, instead, a standard that simply asked whether there was “any evidence” to support the lower court’s findings.²¹ In cases after the amendment, then, the Oregon Court of Appeals reviewed “equitable proceedings” both under an “any evidence” standard and, in cases when it so chose, under a *de novo* standard. This empirical investigation demonstrates that reversal rates did not significantly change under the new standard of review.

This conclusion suggests that, while a change in the standard of review makes a rhetorical difference in the appellate process, it does not drive substantially different outcomes with respect to reversal rates. Rather, a change seems to only reflect a preexisting understanding about the proper role of the reviewing court in comparison to the initial court or agency.

This is not to say that standards of review are irrelevant. They matter rhetorically, and that rhetoric may eventually change the way courts think about the underlying relationships. This Article suggests, however, that (contrary to the traditional understanding) the standard of review does not have an automatic effect on reversal rates. The effect of a standard of review comes not from the mere power of its words, but from how it affects the way appellate courts think about their relationship with the entity whose decision is under review.

Part II of the Article concludes by identifying other moments in recent legal history when the standard of review changed for a particular type of case. From 2000 to 2001, for instance, the United States Supreme Court

21. See *infra* Part II.A (discussing the discretionary *de novo* review in Oregon).

fundamentally altered the standard of review used when reviewing agency decisions that do not speak with the “force of law.”²² And in 1996, the Supreme Court shifted from a *de novo* standard of review to an abuse of discretion standard in the review of decisions that depart from sentences assigned under the federal sentencing guidelines. Future investigations into the effect of these changes on reversal rates can further test the effect of the standard of review on the appellate process and further evaluate the validity of the conclusions developed in this Article.

Part III re-examines the tension between the perspective that differences between standards of review do not make a difference in reversal rates compared to the traditional view that these standards do matter and are critical parts of the appellate process. Ultimately, this Article suggests that this tension can be reconciled by accepting that, in general, standards of review are insignificant to the likelihood of reversal in a particular case, and that the precise phrasing of a standard of review is inconsequential to the outcome of an appeal. But this Article also suggests that the standard of review *is* important as a signaling tool to the reviewing court—as well as to the parties and the decision makers below—indicating a variety of factors that *do* influence how parties approach litigation and appeals, and how they interact with and think about the role that the courts have in resolving decisions in a given case.

This observation leads to two suggestions for entities that seek to define standards of review and for counsel who need to use them. In both instances, this Article suggests that the critical consideration in the definition and application of a standard of review is not the language of the standard itself, but the nature of the relative authority and roles of the reviewing and initial decision-making entities that is reflected in the language of the standard. In other words, it is not what the standard of review *is*, but *why* it is what it is, that most likely affects the outcome of a case. Whether defining the standard of review for the first time, applying an established standard, or changing the standard of review midstream, a particularly careful—and explicit—focus on the assumptions that the standard is making about the relative authority of the parties within the legal system is likely to have more significant results in reversal rates than the mere linguistic exercise of defining, applying, or changing the applicable standards of review.

22. See *infra* Part II.B.1 (describing when the “force of law” analysis was introduced in the United States Supreme Court).

I. STANDARDS OF REVIEW: FUNDAMENTALS

A. *Current Understanding*

This Article divides the standard of review question into two large categories: (1) appellate court review of trial court rulings; and (2) reviewing court examination of administrative agency decisions. While these two categories are, on their faces, rather different, the standards of review used in each context are often compared, and applied with reference, to these different areas of judicial review.²³ As the United States Supreme Court has noted, the history of the recognized standards in these two areas is one where there is a substantial amount of crosspollination between standards for appellate review of judicial decisions and standards for judicial review of administrative decisions.²⁴

23. This analytical crossover between the standards of review can be seen in the early years of both federal administrative agencies and the Federal Rules of Civil Procedure. In 1936, then-Dean of Yale Law School, Charles Clark, explicitly connected the clear error standard of review being proposed for judicially-determined questions of fact under proposed Federal Rule of Civil Procedure 68 and the deferential standard of review of decisions issued by the Federal Radio Commission, predecessor of the Federal Communications Commission (“FCC”), which called for findings of fact, “if supported by substantial evidence, [to] be conclusive unless it shall clearly appear that the findings of the commission are arbitrary or capricious.” Charles E. Clark & Ferdinand F. Stone, *Review of Findings of Fact*, 4 U. CHI. L. REV. 190, 210–11 (1937) (citing 47 U.S.C. § 96(d) (1927) (repealed 1934)). As Clark and Stone note, “it may be suggested that the gap between the two rules seems not overlap.” *Id.* at 211. They later explicitly called for a standard, “perhaps stated in the very words of the amendment to the Radio Act quoted above, to all cases alike,” including judicial facts found in civil cases. *Id.* at 217. In a more recent case, the Supreme Court noted that what appeared to be a conflict between the Federal Circuit’s standard of review for factual determinations made by the Patent and Trademark Office—one explicitly based on the Circuit Court of Appeals’ review of judicial decisions—was effectively no dispute at all, because the Federal Circuit’s “clear error” standard did not look substantially different from the deferential standard under the APA. *Dickinson v. Zurko*, 527 U.S. 150, 163 (1999).

24. “The relevant linguistic conventions were less firmly established before adoption of the APA than they are today. At that time courts sometimes used words such as ‘clearly erroneous’ to describe less strict court/agency review standards.” *Id.* at 156 (citing *Polish Nat’l All. of U.S. v. N.L.R.B.*, 136 F.2d 175, 181 (7th Cir. 1943), *aff’d*, 322 U.S. 643 (1944); *New York Trust Co. v. Sec. & Exch. Comm’n*, 131 F.2d 274, 275 (2d Cir. 1942), *cert. denied*, 318 U.S. 786 (1943); *Hall v. Comm’n*, 128 F.2d 180, 182 (7th Cir. 1942); *First Nat’l Bank of Memphis v. Comm’n*, 125 F.2d 157 (6th Cir. 1942); *N.L.R.B. v. Algoma Plywood & Veneer Co.*, 121 F.2d 602, 606 (7th Cir. 1941)). “Other times they used words such as ‘substantial evidence’ to describe stricter court/court review (including appeals in patent infringement cases challenging *district court* factfinding).” *Id.* (citing *Cornell v. Chase Brass & Copper Co.*, 142 F.2d 157, 160 (2d Cir. 1944); *Dow Chem. Co. v. Halliburton Oil Well Cementing Co.*, 139 F.2d 473, 475 (6th Cir. 1943), *aff’d*, 324 U.S. 320 (1945); *Gordon Form Lathe Co. v. Ford Motor Co.*, 133 F.2d 487, 497 (6th Cir. 1943), *aff’d sub nom. Ford Motor Co. v. Gordon Form Lathe Co.*, 320 U.S. 714 (1943); *Electro Mfg. Co. v. Yellin*, 132 F.2d 979, 981 (7th Cir. 1943); *Ajax Hand Brake Co. v. Superior Hand Brake Co.*, 132 F.2d 606, 609 (7th Cir. 1943); *Galion Iron Works & Mfg. Co. v. Beckwith Mach. Co.*, 105 F.2d 941, 942 (3d Cir. 1939)).

1. Appellate Court Review of Trial Court Decisions

The standard division between appellate court standards of review depends primarily on the relevant question under review. Traditionally, questions of law are reviewed *de novo*,²⁵ questions of fact are reviewed for clear error, and exercises of discretion are reviewed for abuse of discretion.

There are two concerns presented when a new “issue” arises on review in a matter where the standard of review might not be clearly established. First, courts will usually attempt to characterize the nature of the question at issue: Is it a question of law, a question of fact, or an exercise of discretion? Determining the area into which a given appellate issue falls can often be difficult; as Steven Childress has noted, the distinction between these three areas is “somewhat artificial, as law, fact, and discretion may be said to occupy three sides of one strange coin. Or at least in the application of appellate judicial decision making, they often overlap like the circus ringmaster’s three spotlights.”²⁶ Given the uncertain borders between these layers, appellate courts faced with a previously undetermined question will often find themselves characterizing an issue as a “mixed question” of law and fact. A court might review a hybrid issue either *de novo* or deferentially, or it may take a hybrid approach itself (i.e., reviewing certain factual premises deferentially, but then examining the application of law to those facts *de novo*).

De novo review is review without deference. In theory, an appellate court reviewing questions of law makes its own decision about the proper outcome on that question, and does so without reference to, or any necessary reliance on, the trial court’s rulings on the same issue.²⁷ This

25. In Oregon, the term “*de novo*” is reserved for review of facts. Unlike in the federal system, where questions of law are characterized as being reviewed “*de novo*,” in Oregon:

We do not . . . review questions of law “*de novo*.” That term derives from appellate review of equity cases and refers solely to the court’s ability to make its own determination of the facts when it has the authority to “try the cause anew upon the record.” See [OR. REV. STAT. §] 19.125(3) [(2016)]. Although some federal courts misuse the phrase to describe review of issues of law as well as of fact, in Oregon it is clear that it applies only to issues of fact. See O.R.A.P. 5.45 n.1. We assume that plaintiff means to say, as our rules suggest, that we review the trial court’s legal determination for errors of law rather than for abuse of discretion, not that we find facts, something that is beyond our authority in an action at law. See [OR. CONST. art. 7, § 3 (amended 1996); OR. REV. STAT. §] 19.125 [(2016)].

Trabosh v. Washington City, 915 P.2d 1011, 1014 n.6 (Or. 1996).

26. Childress, *supra* note 8, at 1231.

27. See, e.g., *Salve Regina Coll. v. Russell*, 499 U.S. 225, 232–34 (1991) (discussing appellate standards of review for diversity state-law actions); *Appeal de novo*, BLACK’S LAW DICTIONARY

approach is more complicated when it comes to reviewing legal findings of administrative agencies, as is noted below.

Findings of fact made by juries are reviewed differently than findings of fact made by courts. The United States Constitution constrains the power of courts to review jury findings. On appellate review for such a finding, the relevant inquiry is therefore usually phrased as one of law: the jury's finding can be abandoned only if, based on a review of all the evidence, no reasonable jury could have reached the conclusion that it did.²⁸ With respect to judicial findings of fact, Federal Rule of Civil Procedure 52 limits review to reversal only when findings are "clearly erroneous."²⁹

Although the meaning of the phrase "clearly erroneous" is not immediately apparent, certain general principles governing the exercise of the appellate court's power to overturn findings of a district court may be derived from our case. The foremost of these principles, as the Fourth Circuit itself recognized, is that "finding is 'clearly erroneous' when although there is evidence to support it, the reviewing court on the entire evidence is left with the definite and firm conviction that a mistake has been committed." This standard plainly does not entitle a reviewing court to reverse the finding of the trier of fact simply because it is convinced that it would have decided the case differently.³⁰

Finally, discretionary decisions can be reversed only where the trial court has abused its discretion. "There is no one formulation of the abuse of discretion standard. Instead, a highly contextual inquiry varies depending on the setting."³¹ Poor decision-making processes, consideration of factors that should not have been considered, or the failure to consider facts that should have been considered are all circumstances in which an appellate court could reverse a trial court's

112 (9th ed. 2009) ("An appeal in which the appellate court uses the trial court's record but reviews the evidence and law without deference to the trial court's rulings."); 1 STEVEN ALAN CHILDRESS & MARTHA S. DAVIS, *FEDERAL STANDARDS OF REVIEW* § 2.14 (4th ed. 2010) (discussing de novo review).

28. See, e.g., *Hollowell v. Orleans Reg'l Hosp. LLC*, 217 F.3d 379, 385 (5th Cir. 2000) (reasoning that a jury could reasonably find successor liability for a subsequent business when some of the factors determining successor liability were present).

29. See FED. R. CIV. P. 52(A)(6) ("Findings of fact . . . must not be set aside unless clearly erroneous.").

30. *Anderson v. City of Bessemer*, 470 U.S. 564, 573 (1985) (citation omitted). The Seventh Circuit has suggested that reversal for clear error requires a decision to "strike us as more than just maybe or probably wrong; it must . . . strike us as wrong with the force of a five-week-old, unrefrigerated dead fish." *Parts & Elec. Motors, Inc. v. Sterling Elec., Inc.*, 866 F.2d 228, 233 (7th Cir. 1988).

31. Robert Anderson IV, *Law, Fact, and Discretion in the Federal Courts: An Empirical Study*, 2012 UTAH L. REV. 1, 8 (2012).

decision for abuse of discretion.³² In the end, however, consideration under this standard turns on whether the trial court has chosen within a “range of acceptable options”—which affords the trial court substantial discretion.³³

2. Review of Administrative Decisions

While review of trial court and jury decisions dominates the work of federal appellate courts, they are also responsible for reviewing the validity of agency action. Sometimes appellate courts do this directly—as is true for certain regulations promulgated under the Clean Air Act³⁴—and sometimes indirectly—by reviewing, on appeal, a district court’s decision in a run-of-the-mill petition for review under the Administrative Procedure Act (“APA”).

For federal appellate courts,³⁵ the responsibility associated with reviewing administrative decisions is at least as significant a role as that associated with the review of trial court decisions. As with review of trial court decisions, most questions presented for review in administrative law cases fall into three categories: (1) determinations of fact; (2) determinations of law; and (3) decisions on discretionary or policy matters.

Factual determinations by federal agencies are generally set aside only if they are “arbitrary [or] capricious.”³⁶ As the United States Supreme Court has noted, courts traditionally have seen the arbitrary and capricious standard of review as more deferential than the “clear error” standard that courts use to review factual findings by other courts.³⁷ This

32. CHILDRESS & DAVIS, *supra* note 27, § 4.01[3], at 4, 12–15 (discussing how courts apply the abuse of discretion standard).

33. *Id.* § 4.01[1], at 3–4.

34. *See, e.g.*, 42 U.S.C. § 7607(B)(1) (2015) (noting the procedures for judicial review of an “action of the Administrator in promulgating any national primary or secondary ambient air quality standard”).

35. And, to some degree, state courts as well. Although a complete examination into the relationship between state courts and administrative agencies is beyond the scope of this article, the “flavors” of the standards of review of administrative decisions in most states are similar to those in the federal system. There are exceptions and differences in emphasis; in Oregon, for instance, courts review agency decisions for procedural flaws and statutory or constitutional violations, but they do not defer to agency legal determinations. While “substantial evidence” governs judicial review of agency adjudications, OR. REV. STAT. § 183.400 (2016), Oregon courts do not have authority to reverse agency decisions if they are arbitrary or capricious—only if they fall outside the scope of statutory authority. OR. REV. STAT. § 183.482(8)(a)(A).

36. *See* 5 U.S.C. § 706(2)(A) (2015) (providing that agency decisions may be set aside if they are “arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law”).

37. *Dickinson v. Zurko*, 527 U.S. 150, 153 (1999) (“Traditionally, th[e] court/court [clear error] standard of review has been considered somewhat stricter ([i.e.], allowing somewhat closer judicial

is true for factual findings in both informal and formal proceedings. Although facts found in formal agency proceedings are reviewed for “substantial evidence”³⁸—rather than for being arbitrary and capricious—then-Judge Scalia said in 1984 that there is “no substantive difference” between these two tests that review the validity of factual findings.³⁹

Discretionary and policy decisions are also reversed only if they are arbitrary or capricious. This standard is amenable to a wide range of meanings depending on the intensity of this review. On the one hand, this appears to be an extremely deferential standard under which the courts defer to any reasonable and well-informed balance that an agency appears to strike between competing policy objectives. As long as the agency has some kind of rational basis for making its decision, the court will affirm.

On the other hand, courts and academics have also identified a more intensive level of scrutiny under the arbitrary and capricious standard: a so-called “hard look” review. Under this “flavor” of arbitrary and capricious review, the federal courts insist on comprehensive and careful record building by the agency, and engage in an intensive examination of that record to ensure that conclusions necessarily follow from identified facts, that appropriate alternatives were adequately discussed, and that this information is all available in, and articulated on, the agency’s record on review.⁴⁰

Finally, there is the standard of review for questions of law. In the federal system, at least, review of questions of law as decided by agencies can be very different than review of questions of law as decided by lower courts. While appellate review is conducted *de novo* in the latter case, it is highly deferential when it comes to reviewing an agency’s interpretation of law, as long as that interpretation is made in the context of a statement that carries the force of law—such as rules promulgated

review) than the APA’s court/agency standards.” (citing 2 KENNETH CULP DAVIS & RICHARD J. PIERCE, JR., ADMINISTRATIVE LAW TREATISE § 11.2, at 174 (3d ed. 1994)).

38. 5 U.S.C. § 706(2)(E) (2015).

39. *Ass’n of Data Processing Serv. Organizations, Inc. v. Bd. of Governors of the Fed. Reserve Sys.*, 745 F.2d 677, 683–84 (D.C. Cir. 1984). See also EDWARDS & ELLIOT, *supra* note 6, at 182 (characterizing this statement by then-Judge Scalia as “incomplete”).

40. The paradigmatic “hard look” case is *Motor Vehicle Mfrs. Ass’n of U.S., Inc. v. State Farm Mut. Auto. Ins. Co.*, 463 U.S. 29, 43 (1983) (holding that the Court of Appeals erred by not “intensifying the scope of its review based upon its reading of legislative events”). See generally Matthew C. Stephenson, *A Costly Signaling Theory of “Hard Look” Judicial Review*, 58 ADMIN. L. REV. 753 (2006) (discussing problems with “hard look” review administered by a generalist court).

after notice and comment rulemaking—and as long as it is not inconsistent with the plain language of the statute.⁴¹

3. Appellate Court Control over Standards of Review

While the general scope of these standards of review is relatively well-fixed in the firmament of appellate procedure, appellate courts retain a significant amount of discretion with respect to which standard of review applies, as well as the circumstances under which a standard is met in a particular case. Such discretion provides opportunities for appellate courts to drift—whether intentionally or not—from what might be perceived as an “accepted” interpretation of the appropriate intensity of review under a given standard.

Appellate courts have substantial control over three significant areas of standard of review analysis.⁴² First, they can control the characterization of a particular question as one of fact, as one of law, or as a mixed question, thereby leading to a first-order decision regarding the scope of deference in a particular case. If the question on review is one of law, the appellate courts will not be deferential; if the question is one of fact or discretion, the appellate courts will be more deferential; and if the question is a “mixed question of law and fact,” a more sophisticated analysis is required. While a trial court judge certainly has substantial control over how to “set up” this kind of question—by heavily embedding factual findings into any particular determinations, for instance, in the hope that the decision will receive more deference—any significant uncertainty regarding the proper characterization of the issue on review⁴³ remains squarely in the hands of the appellate court.⁴⁴ Depending on the

41. See *Chevron, U.S.A., Inc. v. Nat. Res. Def. Council, Inc.*, 467 U.S. 837, 843–44 (1984) (noting that if a statute is ambiguous, a court may not “substitute its own construction of a statutory provision for a reasonable interpretation made by the administrator of an agency”).

42. This is not to say that appellate courts have plenary authority to determine the scope of review in every case. There are, of course, many areas of the law where the particular standard of review is well-settled. In most cases, this fact will preclude an appellate court from exercising any flexibility on the first point (i.e., whether the question is one of law, fact, or discretion), and limits the scope of flexibility on the second.

43. Given the often shifting lines between issues of fact and issues of law, the scope of issues in which there is uncertainty on this point is quite broad.

44. Consider, for instance, the decision in *Dep’t of Human Servs. v. Three Affiliated Tribes of Fort Berthold Reservation*, 238 P.3d 40, 48–49 (Or. 2010), in which the Oregon Court of Appeals inquired into the proper standard of review for the question whether there was “good cause” to place Native American children outside of their tribe’s preferred placement. While the parties and many prior courts had concluded this was an “abuse of discretion” standard, the Oregon Court of Appeals found otherwise, deeming it a question of law, and reviewable without deference to the lower court conclusions.

motivations of the reviewing court, this is a significant power given that “[l]aw’ and ‘fact’ do not in legal discourse denote pre-existing things; they express policy-grounded legal conclusions.”⁴⁵

Second, once a court has made the first-order characterization of a question as “law” or as “fact,” the appellate court can still modify or carve out the question presented in a way that will dictate the application of a particular standard of review in a particular case. This is especially likely when the relevant issue is a “mixed question of law and fact,” and where the relevant standard has not been clearly determined.⁴⁶ In these circumstances, appellate court control over the phrasing of the question presented, as well as over the ultimate standard used to evaluate that question, is substantial.⁴⁷ In this way, reviewing courts can adjust the phrasing of the standard, or its application in a particular case, to significantly influence the relationship between the appellate court and the decision under review.⁴⁸

Third, once the appellate court determines what the standard of review actually is, the court retains the power to decide whether the standard of review has actually been met. An appellate court might characterize a question on review as one in which it owes deference to the decision maker below, but then proceed to conclude that the decision below should be reversed. While sometimes this is the correct result, there still might be other times in which this outcome is inconsistent with the appropriate standard. The degree to which courts exercise this discretionary power over standards of review in a manner inconsistent with expected outcomes is the focus of much academic skepticism about the predictive

45. Weidner v. Thieret, 866 F.2d 958, 961 (7th Cir. 1989). Judge Posner’s opinion view of this process might be referred to as an “ad-hoc classification decision.” See Martin B. Louis, *Allocating Adjudicative Decision Making Authority Between the Trial and Appellate Levels: A Unified View of the Scope of Review, the Judge/Jury Question, and Procedural Discretion*, 64 N.C. L. REV. 993, 1005 (1986) (noting that an “ad hoc process for classifying ultimate procedural facts achieves different results from the presumptive process employed in the realm of substantive law”).

46. It is in connection with identified “mixed questions” where “many scholars observe the ‘struggle for power’ between trial and appellate courts. Over the years, authors have argued that appellate courts have in fact manipulated the standard of review, often by moving more and more decisions from deferential review into the de novo category.” Anderson, *supra* note 31, at 13.

47. Louis, *supra* note 45, at 1005.

48. In some cases, a given standard of review may even develop formalized “tiers.” Consider, for instance, the distinction between “standard” arbitrary and capricious review under the Federal Administrative Procedure Act and so-called “hard look” version of this review. The stated version of judicial review is the same in both instances—reversal is appropriate only if the agency decision is “arbitrary or capricious”—but in “hard look” cases, the court takes a much deeper and more searching and skeptical look at the validity of the agency’s decision. As I only half-jokingly explain to my administrative law classes, “you know it was ‘hard look’ arbitrary and capricious review when the agency loses.”

power of these standards.⁴⁹

Whether such inconsistency is intentional (i.e., although an appellate court knows that the decision below should objectively be affirmed, it reverses on ideological or personal grounds) or unintentional (i.e., the appellate court mistakenly reverses even though an objective analysis of the case on review would necessarily result in affirmance), this kind of “application error” is very difficult to review or reverse.

The difficulty of review is due not only to the uncertainty associated with the definitions of particular standards, but because cases in which the standard is an issue rarely draw a higher court’s attention. A single error of this type would likely be dismissed at the certiorari stage in the United States Supreme Court as (at worst) an erroneous, fact-bound application of well-settled law that implicates no circuit splits.⁵⁰ Even multiple errors of this kind by a particular circuit court—which might amount to a “drift” in the application of a standard of review over time (e.g., treating arbitrary and capricious review more like clear error review via more regular reversals)—would be unlikely to draw the attention of the United States Supreme Court, which would still see the circuit court as relying on a correct statement of law.⁵¹ A similar point might be made with respect to state high court review of state intermediate appellate courts. The proper application of a given standard of review is difficult to monitor and correct.

The rationales for why appellate courts might exercise this discretion are broad: they might feel personal or political sympathy for a particular party, for instance, or disagree with the underlying policy choices reflected in the cause of action at issue. Many other studies—particularly in the social science literature—have examined the role of political and

49. See, e.g., Anderson, *supra* note 31, at 3–4, 9–10 (discussing the discretionary standards).

50. See S. CT. R. 10 (setting forth standards determining appropriate exercise of certiorari jurisdiction).

51. Persuading the Court to take (and correct) such a drift would be particularly unlikely. First, making the case for a “drift” would certainly require statistical analysis of a particular circuit’s reversal rates, but the Supreme Court has proven to be quite skeptical about the value of statistics about other cases in its review of a single judgment. See, e.g., *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 356 (2011) (explaining that statistical evidence can be insufficient for proving legal theories when the scope of statistical measurement does not match the size of the class or pattern of behavior in controversy); *McCleskey v. Kemp*, 481 U.S. 279, 286 (1987) (attempting to support the claim of racially discriminatory capital sentencing process with a statistical study). Second, the flexibility of, and inherent uncertainty regarding, the line between “affirm” and “reverse” for any standard of review would make it difficult to demonstrate that any given circuit court opinion fell on the wrong side of that line. While a case need not be wrong to justify a grant of certiorari, it certainly helps when making the case for the grant.

policy alignment in how courts make their decisions.⁵²

There are, however, somewhat less nefarious motivations for why appellate judges might skew the results under a particular standard of review. For instance, an appellate court's opinion about the general skill of a particular decision maker below will often inform the outcome of close cases. Such individualized deference is, of course, not a part of the traditional understanding regarding how standards of review are implemented. Nevertheless,⁵³ as D.C. Circuit Judge Patricia Wald once conceded:

[T]here is no question the reviewing judges' perception of the ability of the trial judge counts heavily. If they consider him capable and generally astute, they will try harder to uphold him, especially in close cases. Where errors are asserted, a trial judge of excellent repute will be given the benefit of every doubt.⁵⁴

In *United States v. Boyd*,⁵⁵ for instance, the Seventh Circuit emphasized the experience and wisdom of the trial court judge in noting the deferential nature of the standard of review.⁵⁶

Others have suggested that a typical "signal" regarding a standard of review's reversal rate might be trumped by judicial judgment about the reviewing court's own expertise in the area of law under review. In 2002, for instance, Professor Melissa Hart suggested that the degree of deference by the United States Supreme Court to agency interpretations of law was inversely related to the level of judicial familiarity and

52. For studies that examined this role see generally, CASS R. SUNSTEIN ET AL., *ARE JUDGES POLITICAL? AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY* (2006); Jonathan P. Kestellec, *Panel Composition and Judicial Compliance on the US Courts of Appeals*, 23 J.L. ECON. & ORG. 421 (2007); Jonathan P. Kestellec & Jeffrey R. Lax, *Case Selection and the Study of Judicial Politics*, 5 J. EMPIRICAL LEGAL STUD. 407 (2008); Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717 (1997).

53. Edward H. Cooper, *Civil Rule 52(A): Rationing and Rationalizing the Resources of Appellate Review*, 63 NOTRE DAME L. REV. 645, 655–56 (1988) (noting that although the prospect of this kind of individualized deference is "troubling," it is also likely an accurate description of how appellate courts function).

54. Wald, *supra* note 5, at 1381–82.

55. 55 F.3d 239, 242 (7th Cir. 1995).

56. "[D]uring the trial, Judge Aspen had for months on end listened to witnesses—had heard, had not merely read, their testimony, and had watched them as they gave it A trial judge of long experience, he would have developed a feel for the impact of the witnesses on the jury—and how that impact might have been different had the government played by the rules—that an appellate court, confined to reading the transcript, cannot duplicate. Judge Aspen may have been mistaken; we might suspect that he was mistaken; but unless we are convinced that he was mistaken, we have no warrant to reverse. That is what it means to say that appellate review is deferential. It is not abject, but it is deferential." *Id.* (citations omitted).

comfort with the particular area of law in which the agency was acting.⁵⁷ Noting relatively high reversal rates—and little discussion of deference—in connection with Equal Employment Opportunity Commission legal determinations, Professor Hart hypothesized that this anomalous approach to what should be a deferential standard of review arose out of the United States Supreme Court’s relative familiarity with employment and racial discrimination matters in the context of Title VII and equal protection cases.⁵⁸ That familiarity, she suggested, had a direct consequence on the Court’s willingness to defer to agency interpretations of law that were not otherwise aligned with the Court’s (admittedly informed) view of how the law should operate in that particular area.⁵⁹ In addition to many other factors, then, judicial judgments about their own substantive expertise might trump the signal that Congress or other legislative bodies may send to courts through the standard of review.⁶⁰

For the same reasons that appellate court “drift” in the application of a particular standard would be difficult to catch and resolve,⁶¹ it would be nearly impossible to check an improper application of deference based on an appellate court’s respect for a particular trial court judge, or on its own perception of its expertise. These routes to a “non-standard” standard of review provide another way in which appellate court application of standards of review might shift away from a “standard” application of those principles.

There is, then, substantial control exercised by reviewing courts over the development and application of standards of review in a variety of

57. Melissa J. Hart, *Skepticism and Expertise: The Supreme Court and the EEOC*, 74 *FORDHAM L. REV.* 1937, 1949–58 (2006).

58. *Id.* at 1949–51, 1954–58.

59. *Id.* at 1954–58.

60. The relevant standard of review here is extracted from the United States Supreme Court’s decision in *Chevron, U.S.A., Inc. v. Nat. Res. Def. Council, Inc.*, 467 U.S. 837, 837 (1984). That standard of review is not directly related to legislative action, although it purports to represent a mechanism by which the Court carries legislative intent into effect in the review of agency interpretations of law. *Chevron*, 467 U.S. at 842–45. Under Step 1, a court simply enforces Congressional intent as expressed in clear statutory language, while under Step 2, the court views deference to agency legal interpretations of law as consistent with an implicit Congressional delegation of authority to the agency to announce clarifying statements of law. *Id.*

61. *See* S. CT. R. 10 (emphasizing the compelling character of issues for which certiorari is granted); *McCleskey v. Kemp*, 481 U.S. 279, 319 (1987) (failing to accept statistical evidence in a death penalty case); *see also* *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 356 (2011) (explaining that statistical evidence can be insufficient for proving legal theories when the scope of statistical measurement does not match the size of the class or pattern of behavior in controversy); *but see* *Dukes*, 564 U.S. at 372 n.5 (Ginsburg, J., dissenting) (noting that though the Majority disagrees with the statistical method the court used, “[a]ppellate review is no occasion to disturb a trial court’s handling of factual disputes of this order”).

contexts.⁶² While review of that control might be accomplished by further appeals, such checks are unlikely to work consistently. There are opportunities for reviewing courts to deviate from what might be deemed the “correct” law in this context, and little to be done to correct those errors—unless, of course, specific shifts in the standard of review can influence courts to make such shifts over time. This substantial judicial control over standards of review also suggests that legislative change in the standards—or even change imposed by superior appellate courts—might not have as much influence on judicial decision making as one might assume. Unless courts applying a new standard “buy in” to the proposed change, there are many ways for reviewing courts to stick with preexisting understandings about how deferential they will be.

B. History

The historical evolution of standards of review is a matter that is too involved for this Article to address in great detail, but other scholars have engaged in that analysis to a substantial degree for those who wish to examine the development of these doctrines in greater depth.⁶³ A brief summary is appropriate, however, as it sheds light on the development of the current distinctions between different standards of review, and into the rationale behind the adoption of particular standards.

Like so much that emerged from the development of the common law in the United States, the distinction between law and equity plays a substantial role in the development of today’s appellate standards of review. Historically, “equity review [was] a re-examination of the entire record, on both the facts and the law, while that at law [was] limited to a consideration of the legal errors which may have been committed by the trial court.”⁶⁴ In federal courts, the Judiciary Act of 1789 briefly called for the use of writ of error principles (which did not allow a factual review) in the review of decisions in equitable actions.⁶⁵ Pressure from the “chancery lawyers” resulted in legislation that repealed this combined

62. “The standard of review . . . is potentially a strategic choice for the appellate court and there is the possibility that courts manipulate standards of review.” Anderson, *supra* note 31, at 13.

63. See Erwin N. Griswold & William Mitchell, *The Narrative Record in Federal Equity Appeals*, 42 HARV. L. REV. 483, 486–87 (1928) (examining the evolution of standards of review); see generally CHILDRESS & DAVIS, *supra* note 27 (describing how appellate lawyers can learn how to find, frame, and use the standard of review in practice); Clark & Stone, *supra* note 23 (describing historical significance of the evolution of standards of review).

64. Clark & Stone, *supra* note 23, at 190. The Oregon statute that is the subject of Part III.A. below, reflected this history in that it provided (prior to its 2009 amendment) for de novo review of “equitable action[s].” OR. REV. STAT. § 19.415 (2009).

65. Judiciary Act of 1789, ch. 20, 1 Stat. 73.

standard of review in 1803, however, so appellate courts returned to giving limited deference on review of equitable actions and substantial deference on review of decisions in actions at law.⁶⁶ This distinction remained through 1865, when it was extended to actions at law where juries *and* judges were making findings of fact.⁶⁷ It was only upon the merger of law and equity in the 1938 adoption of the Federal Rules of Civil Procedure that the modern approach to standards of review—based on the nature of the question on review, rather than the legal or equitable nature of the cause of action—became dominant.⁶⁸

This history helps to explain the theoretical development of standards of review. For nearly 150 years, the primary distinction between modes of appellate review turned not on definitions of “fact” or “law,” or on the nature of the relationship between the appellate and trial courts, but rather it turned on the characterization of the cause of action as one at law or one in equity. Essentially, the standard of review turned on the role of the relevant courts in the overall appellate system and the nature of their powers. Once the transformation of the federal judiciary—as well as the judiciaries in nearly all other states—into a merged system was complete,⁶⁹ the standard of review moved from a discussion about how best to characterize the role of particular courts (i.e., as “courts of law,” or “courts of equity”) to a discussion about the nature of particular questions as ones of “law,” of “fact,” or of “discretion.”

Administrative law standards of review for facts tend to parallel civil appellate standards, and are similarly influenced by common law history. But in instances where administrative law differs from civil appellate court-to-court review (i.e., in giving agencies particularly broad “arbitrary and capricious” discretion when it comes to policy choices, or in deferring to reasonable agency interpretations of law that are not otherwise foreclosed by statute), the variations tend to arise primarily out

66. Clark & Stone, *supra* note 23, at 193–96; *see also* Griswold & Mitchell, *supra* note 63, at 487–89 (examining the evolution of legislative efforts at developing the standard of review, specifically noting the history of the Act of March 3, 1803).

67. Clark & Stone, *supra* note 23, at 197–99; *see also* Griswold & Mitchell, *supra* note 63, at 513 (noting how district courts in equity cases and in actions at law when the jury was waived were required to make findings of facts).

68. Another good summary of the history of appeals in equity practice can be found in Maury Holland & Robert L. Sokol, *A Statutory Antique: Rethinking Oregon’s Civil ‘De Novo’ Appellate Review*, 66 OR. ST. B. BULL. 19 (May 2006).

69. The formal merger of law and equity in the federal system came in 1938; in states, some preceded the federal system with a formal merger, while others did not take formal steps to complete the merger until well into the twentieth century. As Part III.C. notes below, many state appellate systems still retain vestigial parts of the distinctions between the two systems.

of a motivation that recognizes that administrative agencies are different institutions from trial courts, and they play different roles in the legal system. In both civil and administrative systems of review, then, the relevant standard of review is rooted in historical and current understandings about the relative role of the deciding and reviewing institutions within the overall system.

C. Purposes

A historical perspective, then, emphasizes the importance of institutional roles in the development and maintenance of the current standards of review. Under this perspective, a standard of review is fundamentally a mechanism that describes—and, perhaps, mediates⁷⁰—the relationship between the reviewing court and the decision maker under review. It represents a tangible means of measuring the relative authority and competence of the two decision-making bodies—the reviewing court and the entity (usually the trial court or an agency) that issued the decision under review—that are a necessary part of any review process.⁷¹ As one commentator suggested, “[t]he organization of the federal judiciary is premised on the division of labor between trial and appellate courts, the boundaries of which are delineated by the standards of review.”⁷² The same might also be said of state judiciaries.

Other factors explain and influence standards of review, of course. “There are, to be sure, many additional considerations governing the allocation of responsibilities between the trial and appellate levels, including judicial economy, the enhancement of public regard for the functioning of the judicial system, and the need for both consistent decisions and clear rules.” But in the end, however, “[i]t is clear . . . that institutional competence is the dominant consideration” associated with the creation and definition of standards of review.⁷³

Implicit in this explanation for the standard of review is the notion that when one court has more “expertise” in playing a particular role within the appellate system, it will issue “better” decisions in that particular area than another entity might. For example, appellate courts typically give

70. This is the assumption, at least, of those who emphasize the importance of standards of review in appellate rhetoric.

71. See, e.g., Anderson, *supra* note 31, at 44–48 (explaining that institutional competence—“the idea that by their situation or their experience[,] the trial judges find facts better than appellate judges”—is the most common reason for deferential review).

72. *Id.* at 6.

73. Chad M. Oldfather, *Appellate Courts, Historical Facts, and the Civil-Criminal Distinction*, 57 VAND. L. REV. 437, 445 (2004).

less deference to trial courts in questions of law because they are viewed as having particular expertise in making—and, in fact, they are designed to ensure—good legal decisions. By contrast, some view trial courts to be better than appellate courts at gathering, assessing, and evaluating facts, even if the trial court’s evidence is purely documentary. Therefore, appellate review of facts is highly deferential to trial court findings (or jury findings).⁷⁴

Given this background, one might think of standards of review as describing—rather than prescribing—the nature of the authority and competencies of the institutions involved in appellate review. Under this theory regarding the function of standards of review, one would expect institutions to be relatively resistant to a change in a standard of review, unless the change simultaneously occurred with a shift in the underlying understanding about that relationship. Similarly, one would expect that data regarding reversal rates would reflect the importance of standards between different types of cases because the standards track established understandings. Ultimately, however, one might expect this effect to be rather weak because appellate court control over the application of a standard, and the influence of the court’s perception of its role in the process, is likely to be more significant than the “forcing” effects associated with the language of the standard itself.

D. Does a Different Standard Affect Outcomes?

1. Different Standards Affect Outcomes (Circumstantial Evidence)

Circumstantial evidence suggests that the standard of review matters significantly. Generally speaking, of course, the standard itself is phrased as a decision rule; the reviewing court may not reverse unless the standard of review has been met. Not surprisingly, an effective appellate attorney therefore spends a substantial amount of time crafting briefs and preparing his or her oral argument with the standard of review in mind. Appellate judges add to the perceived importance of the standard because their decisions can dwell extensively on the standard and whether it is met in a particular case. Given the focus on the standard in each appellate case, it is not at all unreasonable to assume that a given standard makes a difference to the outcome of the issue on review.

74. “[T]he degree of deference accorded the trial judges is designed to promote the better functioning of the judicial process. In some circumstances, it has been believed, the trial judge is in a better position to make a correct decision than are appellate judges.” MEADOR ET AL., *supra* note 1, at 222; *but cf.* Oldfather, *supra* note 73, at 440 (arguing that sometimes appellate courts are better than district courts in evaluating factual findings).

This assumption is reinforced by the obligation of appellate counsel to include a statement regarding the standard of review applicable to their particular case in their opening briefs. In the federal system, the Federal Rules of Appellate Procedure have included such a requirement since 1993, and individual circuits asked for such a statement for many years prior to that.⁷⁵ State courts generally impose similar requirements.⁷⁶ The mandatory nature of these statements emphasizes the importance of these standards in the minds of practitioners and judges.

Similarly, courts regularly spend a significant amount of time in cases evaluating the standard of review. Specifically, cases in which no standard is currently established require parties and the courts to expend significant resources in evaluating and deciding on the relevant standard.

The language of the standards certainly suggests their value. And this seems to make sense; as a rhetorical matter, there is a substantial difference between a case where an appellate court can reverse only when the decision below is “arbitrary and capricious,” and a case where an appellate court’s review is *de novo*. It is not particularly difficult to theoretically imagine a case in which the court believes the decision below to be incorrect, but perhaps not so incorrect that it amounts to an arbitrary decision worthy of reversal. As courts repeatedly remind practitioners in discussing deferential standards of review, a court using such a standard may not reverse merely because the court believes the decision below to be incorrect; rather, the court reverses only when the decision below cannot be *supported*. Ultimately, the focus of courts, practitioners, and observers on the standard of review provides strong anecdotal evidence that they make a difference.

75. FED. R. APP. P. 28(a)(8)(B) (requiring “for each issue, a concise statement of the applicable standard of review”). The Advisory Committee notes accompanying the adoption of the rule noted that “five circuits currently require these statements.” *Id.* (noting the 1993 Notes of Advisory Committee on Rules).

76. OR. R. APP. P. 5.45(5) (“Under the subheading ‘Standard of Review,’ each assignment of error shall identify the applicable standard or standards of review, supported by citation to the statute, case law, or other legal authority for each standard of review.”). The Oregon rules also helpfully set forth a range of options so that there is no mistaking to what the rule refers. OR. R. APP. P. 5.45(5) n. 2 (“Standards of review include but are not limited to *de novo* and substantive review for factual issues, errors of law and abuse of discretion for legal issues, and special statutory standards of review.”). See also ARIZ. R. CIV. APP. P. 13(a)(7)(B) (requiring statement of “the applicable standard of appellate review with citation to supporting legal authority”); COLO. APP. R. 28(k) (requiring, “[f]or each issue raised on appeal, . . . under a separate heading placed before discussion of the issue: (1) a concise statement of the applicable standard of appellate review with citation to authority.”); FLA. R. APP. P. 9.210(b)(5) (briefs shall contain “Argument with regard to each issue, with citation to appropriate authorities, and including the applicable appellate standard of review”). Not all courts explicitly require such a statement, apparently assuming that counsel will rely on their training and writing skills. *E.g.*, CAL. R. APP. PRO.; WISC. R. APP. PRO.

Furthermore, despite the uncertainty in some areas of the law, there are particular questions for which the relevant appellate standard is well-established. Even when there is no clear guidance on the standard of review, at least some commentators have dubbed the “basic test for whether deferential review is appropriate” (i.e., the determination of whether a question of law or a question of fact is under review) to be “remarkably functional” in its focus on the simple questions of competence of the relevant decision-making body.⁷⁷ Perceiving the straightforward nature of this baseline test, many judges continue to argue for the importance of the traditional emphasis on the choice between different standards.⁷⁸ In response to academic skepticism regarding the importance of standards of review, for instance, one judge on the U.S. Court of Appeals for the D.C. Circuit characterized that skepticism as “misguided” and “cynical.”⁷⁹ As the next section demonstrates, this perception is reflected, at least in part, by the empirical literature that finds an outcome-oriented effect of standards of review.

2. Different Standards Affect Outcomes (Empirical Findings)

Empirical data seem to provide the best opportunity to determine whether a standard of review matters to appellate outcomes. Some empirical studies suggest that standards of review make a difference, at least in certain circumstances.⁸⁰ This evidence is limited, however; as one 2012 study noted, the question of whether the reversal rate correlates with the standard of review had not been empirically studied until recently.⁸¹

A 2010 empirical investigation conducted by Corey Rayburn Yung concluded that standards of review *did* matter.⁸² Based on his examination of federal appellate cases issued in 2008,⁸³ Yung found that appeals using a *de novo* standard of review resulted in reversal approximately 25.5 percent of the time, appeals using an abuse of discretion standard of review led to reversal about 15 percent of the time,

77. Anderson, *supra* note 31, at 15.

78. See *supra* text accompanying notes 2–6 (describing the traditional distinction between standards).

79. EDWARDS & ELLIOTT, *supra* note 6, at v.

80. See *infra* text accompanying notes 82–93 (describing such empirical studies).

81. Anderson, *supra* note 31, at 10–11 (citing Yung, *supra* note 15, at 1136). While both Yung and Anderson’s studies of this question were driven primarily by efforts to characterize the effect of ideological leanings on case outcomes, they each examined whether the standard of review made a difference in reversal rates in the course of conducting that investigation.

82. Yung, *supra* note 15, at 1159.

83. *Id.* at 1155–56.

and cases involving clear error review led to a 12 percent rate of reversal.⁸⁴ Despite the conventional wisdom of social science that “law has almost no influence on” judges,⁸⁵ Yung found that the spectrum of review seemed to have predictive value, and concluded that standards of review *did* influence the reversal rate.⁸⁶

In Yung’s study, however, there were some potentially confounding factors. Most concerning, perhaps, was that because of the search algorithm used, the dataset included only those cases in which the appellate court “used language relevant to a standard of review.”⁸⁷ This would seem to be a particularly difficult selection problem if one is attempting to determine whether standards of review are appropriately applied across *all* appellate cases because one would expect the standards to be most clearly articulated in cases where the court wanted to hew the standard itself—whether that involved reversing or affirming in the cases where the standard mattered. It does not seem surprising, then, that the effect of the standard seemed particularly strong in Yung’s study.

Yung’s result was replicated, with some subtle differences, in Robert Anderson’s 2012 study. In that study, Anderson concluded that, “deferential standards of review appear to considerably decrease the probability of outright reversal.”⁸⁸ Curiously, however, they did not increase the probability of affirmance.⁸⁹ Anderson suggested that this might be due to “mixed” outcomes (i.e., partial affirmance and partial reversal), which seemed to increase with more deferential standards of review.⁹⁰ So while there was some relationship to reversal, his data did “not show the clear relationship between the affirm/reverse decision expected by the legal model [which assumes that judges attempt to comply with legal standards, without reference to outside considerations], but instead show[ed] some surprisingly nuanced relationships among the standards of review and the outcome of the appeal.”⁹¹

84. *Id.* at 1161 fig.1.

85. *Id.* at 1161 n.217 (citing Howard Gillman, *What’s Law Got to Do With It? Judicial Behavioralists Test the “Legal Model” of Judicial Decision Making*, 26 LAW & SOC. INQUIRY 465, 466 (2001)).

86. Yung, *supra* note 15, at 1160–61.

87. *Id.* at 1195; *see also id.* at 1156 n.188 (describing the Lexis search that returned examined cases).

88. Anderson, *supra* note 31, at 5.

89. *Id.* at 23–26.

90. *Id.* at 25.

91. *Id.* at 24. A skeptic might argue that “surprisingly nuanced” results associated with what should be a straightforward spectrum of standards (i.e., the hypothesis that more deference leads to

One of the central findings of Anderson's study is that standards of review provide a mechanism through which differences in standards of review might alter rates of reversal. It suggests that the fact of deference—whether accurate or not—might incentivize trial courts to write their decisions in a way that grounds their opinions in deferentially reviewed rationales.⁹² The idea that trial courts act strategically to modify their decisions and take advantage of the deference that appellate courts owe is not unusual,⁹³ but Anderson is not able to directly test whether this supposed manipulation of the standard of review works (i.e., whether it leads to different reversal rates at the reviewing court level).

In a 2011 study by Fischman and Schanzenbach, their data suggested, but did not affirmatively demonstrate, that appellate courts are “constrained by standards of review.”⁹⁴ This study focused on whether district court judges were influenced by changes in the standard of review associated with departures from the federal sentencing guidelines.⁹⁵ Fischman and Schanzenbach concluded that district courts *were* affected by the change to a more deferential standard of review.⁹⁶ They further hypothesized that this effect would take place only if appellate court judges were constrained by review standards, and suggested that their work “provide[s] indirect evidence that circuit court panels are also constrained by standards of review.”⁹⁷

On the other hand, the evidence gathered by Fischman and Schanzenbach could be explained differently. It seems equally likely that district courts could change their behavior based on the *perception* that appellate courts were changing theirs, without actually being able to predict (or determine after appeal) whether appellate courts were actually changing their behavior given the change in the standard of review. As long as parties and the trial courts believed that appellate courts were going to be lenient, perhaps, they decided to depart more often. Thus, Fischman and Schanzenbach might be correct that trial courts are “meaningfully constrained by the *prospect* of appellate reversal,”⁹⁸ but that might well be true even if trial courts had no idea whether the

more affirmance) are evidence that the hypothesis is incorrect.

92. *Id.* at 9 (examining several models for how standards of review work in the federal system).

93. *See infra* text accompanying notes 94–97 (describing empirical studies on district court behavior and “constrained” standards of review).

94. Joshua B. Fischman & Max M. Schanzenbach, *Do Standards of Review Matter? The Case of Federal Criminal Sentencing*, 40 J. LEGAL STUD. 405, 406 (2011).

95. *Id.* at 406. One of the changes to the standard that is encompassed within the study is the change that is the topic of Part II.B., below.

96. *Id.* at 431.

97. *Id.* at 406.

98. *Id.* at 405.

likelihood of appellate reversal was going to change. It would seem very difficult for trial courts to accurately predict whether a certain departure might be subject to reversal in a given case—let alone whether the statistical evidence might suggest such an outcome. For a trial court judge, the outcome of an appeal is not a statistical reversal rate—it is either “affirmed” or “reversed”—and it seems unlikely that the trial court judge would have the necessary data with which to evaluate statistical changes in the risk of such a result. Thus, while Fischman and Schanzenbach may be right that reversal rates differ with a changed standard of review, that conclusion is not necessarily mandated by their findings.

Finally, Yung’s and Anderson’s studies,⁹⁹ at least, confirmed earlier work conducted by Frank Cross on published federal appellate court decisions from 1928 to 1988.¹⁰⁰ In his work, Cross found that standards of review seemed to influence outcomes in certain civil appeals, and that the spectrum of standards correlated with reversal rates in cases involving review of decisions by federal administrative agencies.¹⁰¹

Collectively, the above studies suggest that the standard of review does make a difference to appellate practice. Yet, other information suggests the contrary.¹⁰²

3. Different Standards Do Not Affect Outcomes (Circumstantial Evidence)

For every statement by a practitioner that standards of review make all the difference in the world, there are equal numbers of statements

99. See Anderson, *supra* note 31, at 50 (examining in both studies the effects of standards of review in reversal rates); Yung, *supra* note 15, at 1159–61 (examining Cross’ empirical studies regarding the connection between standards of review and reversal rates).

100. See generally Frank B. Cross, *Decisionmaking in the U.S. Circuit Courts of Appeals*, 91 CALIF. L. REV. 1457 (2003) (examining determinants of decision making in the United States Circuit Courts of Appeals by conducting an empirical study of determinants, using a database that includes thousands of decisions).

101. See *id.* at 1501–03 (finding inverse correlation between reversal rates and magnitude of anticipated deference to trial court decisions (like summary judgment, review of jury trial decisions, and review of denials of post-trial motions)). In the case of agencies, the affirmance rates across several standards of review varied by less than 10 percent, ranging from 67 percent affirmance in de novo review cases to 75 percent affirmance in reviews deploying the arbitrary and capricious test. While the order of reversal rates is what one would expect (the more substantial the deference, the more likely the court was to affirm), the spread seems narrower than one might hypothesize based on judicial descriptions of the strength of these standards. Cf. Verkuil, *supra* note 16, at 689 (hypothesizing de novo affirmance rates of 50 percent and arbitrary and capricious affirmance rates of 85 to 90 percent, but not discussing whether the small differences are statistically significant).

102. See *infra* Parts I.D.3–4 (noting circumstantial evidence and empirical findings that support the idea that different standards do not affect outcomes).

suggesting the contrary. Many of these statements come from academics evaluating the importance of standards of review. For example, Gellhorn and Robinson famously stated that in the context of judicial review of administrative decisions, “the rules governing judicial review have no more substance at the core than a seedless grape.”¹⁰³ Nor is it particularly unusual for courts to challenge practitioners with the (apparently difficult) task of finding decisions where differences in the standard would really matter,¹⁰⁴ or even for courts to suggest that certain differences in the characterization of the standard of review are likely to be irrelevant.¹⁰⁵ Judge Richard Posner suggested that there are effectively just two standards—deferential, and not—and that other standards with subtle differences are not “actually capable of being drawn in the practice of appellate review.”¹⁰⁶ One commentator suggested, in evaluating the “clearly erroneous” language of Federal Rule of Civil Procedure 52, that the phrase does not have an “intrinsic meaning. It is elastic, capacious, malleable, and above all variable. Because it means nothing, it can mean anything and everything that it ought to mean. It cannot be defined, unless the definition might enumerate a nearly infinite number of shadings along the spectrum of working review standards.”¹⁰⁷

Even more fundamentally, some have pointed out that the “distinction between the terms ‘law’ and ‘fact’ is malleable at best, and the malleability is not entirely unwelcome to the judges who wield the terms.”¹⁰⁸ Ultimately, some suggest, “it may be that the standards governing appellate review in fact provide very little governance at all.”¹⁰⁹

103. Ernest Gellhorn & Glen O. Robinson, *Perspectives on Administrative Law*, 75 COLUM. L. REV. 771, 780 (1975).

104. See, e.g., *Rose v. State*, No. 59141, 2013 WL 3297397, at *3 (Nev. May 13, 2013) (noting that counsel “offers no cogent argument as to how a different standard of review on appeal would have affected the outcome of either the trial or the appeal”).

105. See, e.g., Sch. Dist. of Wisconsin Dells v. Z.S., ex rel. Littlegeorge, 295 F.3d 671, 676 (7th Cir. 2002) (noting Judge Posner’s opinion that even if the lower court applied the wrong standard of review in the case at bar, it was harmless); see also *Ass’n of Data Processing v. Bd. of Governors*, 745 F.2d 677, 686 (D.C. Cir. 1984) (noting Justice Scalia’s opinion that the “‘substantial evidence’ requirement . . . demands a quantum of factual support no different from that demanded by the substantial evidence provision . . . which is in turn no different from that demanded by the arbitrary and capricious standard”); but see EDWARDS & ELLIOTT, *supra* note 6, at 182–83 (challenging Justice Scalia).

106. *United States v. Boyd*, 55 F.3d 239, 242 (7th Cir. 1995).

107. Cooper, *supra* note 53, at 645.

108. Anderson, *supra* note 31, at 12; see also *Miller v. Fenton*, 474 U.S. 104, 113 (1985) (“[T]he appropriate methodology for distinguishing questions of fact from questions of law has been, to say the least, elusive.”).

109. Oldfather, *supra* note 73, at 439.

The wide range of discretion that appellate courts wield in choosing and implementing particular standards¹¹⁰ provides reviewing courts with substantial opportunities to deviate from the typical “textbook description of the adjudicatory process”¹¹¹ under which they impartially strive to implement established rules of legal reasoning to evaluate the validity of the lower court’s decision. If that is true, one would expect the availability of this discretion to confirm the skepticism of observers by showing itself in empirical studies on the effect that standards of review have on outcomes. As described in the next section, some studies do, in fact, conclude that reversal rates are not affected by changes in the standard of review.

4. Different Standards Do Not Affect Outcomes (Empirical Findings)

Uncertainty in the findings of some of the above studies—specifically in Anderson’s study¹¹²—suggests that there is not a straightforward relationship between the standard of review and the outcomes. Perhaps the most important study that directly contradicts such a proposition is the 2002 study by Paul Verkuil.¹¹³ In his study, Verkuil examined the reversal rates on review of three different types of agency decisions: Social Security Administration (“SSA”) disability determinations, Veterans’ Disability determinations, and Freedom of Information Act (“FOIA”) decisions.¹¹⁴ The decisions were all reviewed under very different standards: (1) by statute, SSA disability decisions are reversed on factual grounds only if they are not supported by “substantial evidence”¹¹⁵; (2) fact finding in Veterans’ disability determinations is reviewed for clear error¹¹⁶; and (3) FOIA decisions are, by law, reviewed *de novo*.¹¹⁷ Based on these standards, Verkuil hypothesized that the SSA decisions would be reversed rarely (i.e., at a 15 to 25 percent rate), and FOIA decisions relatively often (surviving review only about 50 percent of the time), with the reversal rate for Veterans’ Disability determinations

110. See *supra* Part I.B. (discussing the discretion appellate courts have in choosing the standard of review).

111. Anderson, *supra* note 31, at 2.

112. *Id.* at 24–25.

113. See generally Verkuil, *supra* note 16 (arguing that relevant standards of review do not reflect reversal rates).

114. *Id.* at 702–18.

115. 42 U.S.C. § 405(g) (2015); see *Masterson v. Barnhart*, 309 F.3d 267, 272 (5th Cir. 2002) (reviewing the denial of benefits only to ascertain whether substantial evidence supports the final decision and whether the Commissioner used the proper legal standards to evaluate the evidence).

116. 38 U.S.C. § 7261(a)(4) (2015).

117. 5 U.S.C. § 552(a)(4)(B) (2015).

somewhere between the two (around 20 to 30 percent).¹¹⁸

After reviewing years of decisions, however, Verkuil found nearly the opposite.¹¹⁹ FOIA decisions were affirmed nearly 90 percent of the time, despite being subject to *de novo* review.¹²⁰ He also found that SSA disability decisions, by contrast, were reversed nearly half the time.¹²¹ And, further, Veterans' disability decisions were the only decisions found close to meeting Verkuil's hypothesized reversal rates.¹²² Verkuil concluded that the hypothetical reversal rates were nearly useless in predicting relative outcomes between these different areas of the law, and that the hypothesized reversal rates "dramatically" failed to materialize.¹²³

Part of the difficulty with Verkuil's examination is that the different reversal rates arise out of altogether different decisions made by entirely different decision-making bodies. While it certainly demonstrates the fallacy of the proposition that standards of review are consistent and meaningful across all comparisons, his study also might suggest that there are simply strong differences in the way that courts approach the underlying substantive area of law, or in how courts perceive the quality of agencies, or at least the quality of agency decision making within particular substantive areas.¹²⁴

Other studies support the conclusion that standards of review do not necessarily predict outcomes—or at least that they do not *completely* do so. For instance, while judicial findings of fact in civil and criminal courts are both subject to a clear error standard of review, "mounting empirical evidence . . . suggests that appellate courts routinely reexamine the evidence supporting a trial verdict in civil cases, but almost never do so in criminal cases."¹²⁵ The test for overturning a jury verdict is no different in the two systems; therefore, the difference in reversal rates demonstrates that the standard of review is being modified by other

118. Verkuil, *supra* note 16, at 689.

119. *Id.* at 719.

120. *Id.*

121. *Id.*

122. *Id.*

123. *Id.*

124. His work suggests that there would be great value in examining reversal rates among different types of decisions made by the same agency to determine whether judicial skepticism (or comfort) regarding the validity of agency decisions changed as the nature of the substantive areas changed (i.e., perhaps SSA's decisions in a different area of benefits varies substantially from the reversal rates for disability determinations), or if the reversal rates work of particular agencies remained consistent to the agencies.

125. Oldfather, *supra* note 73, at 439 (collecting cases). As Oldfather notes, this should be an "astonishing revelation" given the constitutionally-driven understanding of the importance of skepticism when it comes to criminal convictions. *Id.* at 441–42.

considerations.

At the very least, this collected empirical evidence suggests that standards of review might not always do a particularly good job of predicting reversal rates. If that is true—at least sometimes—and if the standards are so malleable that a particular standard (even if it is clearly the one that should be applicable) might not be applied in a predictable way, then the underlying validity of how standards of review affect outcomes should be questioned.

5. Additional Evidence Regarding the Effect and Nature of Appellate Standards of Review

One final set of studies is worth mentioning. In cooperation with several other academics, Ted Eisenberg examined appellate reversal rates in both federal and state courts in an effort to ascertain whether there is some kind of systemic bias among appellate court judges.¹²⁶ In both cases, Eisenberg concluded that there was a consistent bias against plaintiffs at the appellate level.¹²⁷

That conclusion is only so useful in evaluating the question at issue in this Article: whether a change in the standard of review changes the rate of reversal. While ideology or general predilections for one party or another might drive absolute differences in reversal rates, the primary focus of this Article is on the relative difference in reversal rates as courts move from one standard of review to another. Even if courts were to abandon standards of review altogether and appellate courts simply reversed if a trial court was “wrong,” the party biases identified by Eisenberg would presumably still exist in differential reversal rates. But the critical issue in this Article is whether the difference in standards of review *results* in different reversal rates.

The Eisenberg studies provide one interesting and useful point for purposes of this Article: they reveal that appellate reversal rates are affected, at least in part, by appellate court perceptions of the work of trial courts and advocates within them. The results of the Eisenberg studies “suggest that appellate court attitudes and assumptions about trial courts likely shape the observed pattern of appeals’ outcomes even after

126. Theodore Eisenberg & Michael Heise, *Plaintiphobia in State Courts? An Empirical Study of State Court Trials on Appeal*, 38 J. LEGAL STUD. 121, 125 (2009); see Kevin Clermont & Theodore Eisenberg, *Plaintiphobia in the Appellate Courts: Civil Rights Really Do Differ from Negotiable Instruments*, 2002 U. ILL. L. REV. 947, 949 (2002) (contrasting trial court bias and appellate court bias and finding that “[a]ppellate courts are indeed more favorable to defendants than are trial judges and juries”).

127. Eisenberg & Heise, *supra* note 126, at 124; Clermont & Eisenberg, *supra* note 126, at 949.

accounting for the selection of cases to appeal.”¹²⁸ If one assumes that plaintiffs and defendants face similar standards of review,¹²⁹ the Eisenberg studies suggest that the reversal rate under the same standard is nevertheless different, and that the standard of review, in itself, is a poor predictor of reversal rates.¹³⁰ Those studies do not, however, perceive the difference to be rooted in ideology, invidious bias, or general opposition to “plaintiffs.” Their conclusion is that this “asymmetry” problem is rooted in the appellate courts’ (mis?)perception of excess trial court (and especially excess jury) tolerance for plaintiffs at the trial court level.¹³¹

There are, then, hints in the above information that differences in reversal rates are driven not by a strict application of the standards, but by a direct perception of relative institutional competence. One should not dismiss the possibility that scope of any difference in reversal rates for different standards of review expands or contracts based on other factors—such as judicial politics, background, training, or the general nature of a given case.¹³² This Article, however, contends that the best way to test the underlying questions—do different standards of review affect outcomes, and if so, how?—is by examining situations in which similar (or identical) decision makers, facing a similar (or identical) mix of cases, are suddenly faced with a change in the stated standard of review. Part II is directed to that task.

II. CHANGING STANDARDS OF REVIEW: THREE EXAMPLES

A. *Discretionary De Novo Review in Equitable Actions or Proceedings in Oregon, 2005–2014*

Prior to 2009, the Oregon Court of Appeals was required, pursuant to Oregon Revised Statutes (“ORS”) 19.415(3), “[u]pon an appeal from a judgment in an equitable proceeding,” to try the case “anew upon the

128. Eisenberg & Heise, *supra* note 126, at 124; see Clermont & Eisenberg, *supra* note 126, at 949 (noting that appellate judges lean in favor of the defendant potentially as a result of their misperceptions about trial courts’ reputation of being pro-plaintiff).

129. This may not be the case; given the different postures that the parties face, as well as the plaintiffs’ burden of proof, their analysis regarding whether to appeal from adverse rulings in particular areas may change the balance of issues—and therefore the balance of standards—that the parties bring on appeal.

130. Eisenberg & Heise, *supra* note 126, at 123–24.

131. *Id.* at 144–45.

132. There are hints of these effects in some of the above-referenced articles, many of which seek to evaluate the effect (for instance) of ideology on reversal rates or on trial court decisions. *E.g.*, Anderson, *supra* note 31, at 10–18.

record.”¹³³ As a 2006 commentary on this provision noted, its roots could be found in the historical lines between courts of law and courts of equity.¹³⁴ As an historical matter, appeals in equity cases were viewed as “more akin to retrials, whereby they would be ‘tried anew upon and regard to all questions both of law and fact presented by the transcript,’”¹³⁵ than to appeals of common law cases, in which appellate review of factual findings was, and remains, severely circumscribed by both federal and state Constitutional limitations.¹³⁶ This statute, then, was viewed as establishing—or, perhaps more accurately, retaining—“de novo appellate review of fact findings in equity cases.”¹³⁷

As was true in many jurisdictions with this mode of appellate review in equity cases, the Oregon state courts developed an approach to the review of facts in equity cases that was nevertheless quite deferential. As Charles Clark noted in 1936, many courts then adhered to the old line of demarcation between no review of fact in cases at law, and de novo review in cases at equity, but construed “away its force by extensive application of the presumption in favor of the trial court’s finding.”¹³⁸ This was the case in Oregon; even in cases that were understood to call for the review of facts “anew upon the record,” the appellate court would defer to a trial court’s findings of fact, reasoning that the trial court’s ability to observe witnesses and hear testimony “should be given great weight.”¹³⁹ Thus, despite the statutory call for de novo review, the courts gave significant weight to their perception of the institutional competencies of the two courts.

The nature of appellate review under this Oregon statute became so deferential that at least some courts occasionally characterized their de novo review of facts in these cases through the language of entirely

133. Prior to 2009’s amendment, OR. REV. STAT. § 19.415(3) (2005) provided: “(3) Upon appeal from a judgment in an equitable proceeding the Court of Appeals shall try the cause anew upon the record.”

134. Holland & Sokol, *supra* note 68, at 19.

135. *Id.* at 19 n.4 (citing *Heatherly v. Hadley*, 4 Or. 1, 9 (1869)).

136. See OR. CONST. art. VII (amended), § 3 (“[N]o fact tried by a jury shall be otherwise re-examined in any court of this state, unless the court can affirmatively say there is no evidence to support the verdict”); see also U.S. CONST. amend. VII (“[N]o fact tried by a jury shall be otherwise reexamined in any Court of the United States, than according to the rules of the common law.”).

137. Holland & Sokol, *supra* note 68, at 19 n.4 (emphasis omitted).

138. Clark & Stone, *supra* note 23, at 215.

139. *Paulson v. Paulson*, 404 P.2d 199, 202 (Or. 1965); see *Cline v. Larson*, 418 383 P.2d 74, 89 (Or. 1963) (noting that in equity cases, when “the inferences and innuendoes to be drawn from the testimony are several, great reliance is to be placed upon the findings of the trial judge”) (citing *Clauder v. Morser*, 282 P.2d 352, 358 (Or. 1955)).

different standards of review. In 1998, for instance, the Oregon Court of Appeals reversed findings of fact from a circuit court decision in an equitable matter; in doing so, it stated “[o]n de novo review, ORS 19.415(3)[,] we conclude that there is clear and convincing evidence that plaintiff satisfied all the requisite elements of her adverse possession claim and that there is no basis for estopping her from asserting it. We therefore reverse.”¹⁴⁰

The original version of ORS 19.415(3), then, was viewed as one of “Oregon’s most anachronistic statutes, a veritable museum piece of an enactment,” and the courts apparently viewed the teachings of the statute with skepticism.¹⁴¹ In 2009, the Judicial Department asked the legislature to amend the statute “to conserve court resources.”¹⁴² As is suggested by the Judicial department’s justification for the change, the courts believed that the change would not alter the rate of reversal—despite the shift from de novo to “clear error” review.¹⁴³ Rather, the courts believed that this change would make the appellate process easier by limiting the scope of factual recitations in the briefs and the degree to which the Courts believed that they had to recite facts in their opinions.¹⁴⁴ Based on this request, the Oregon Legislative Assembly amended ORS 19.415(3), thus giving the Court discretion to review cases de novo, but not *requiring* it to do so.¹⁴⁵ The statute retained the de novo requirement for cases

140. *Meier v. Rieger*, 954 P.2d 786, 788 (Or. Ct. App. 1998), *abrogated on other grounds by Hoffman v. Freeman Land and Timber, LLC*, 994 P.2d 106 (Or. 1999) (emphasis omitted).

141. *Holland & Sokol*, *supra* note 68, at 19.

142. *See* S. 262, 75th Legis. Assemb. Reg. Sess. (Or. 2009) (“[Senate Bill] 262A was introduced on behalf of the Judicial Department in an effort to streamline court functions and reduce costs. Oregon is one of the few states that provides universal de novo—that is, trying the case anew on the record—review in equity cases. [Senate Bill] 262A requires de novo review in only a few types of cases in order to conserve court resources.” (emphasis omitted)).

143. *Id.*

144. Notes of Court of Appeals Chief Judge D. Brewer, Remarks at Continuing Legal Education Program (Jan. 22, 2010) [hereinafter Brewer, CLE Remarks] (on file with the author).

145. Or. Rev. Stat. §19.415(3) (2009). After the amendment, the statute provided:

(3) Upon an appeal in an equitable action or proceeding, review by the Court of Appeals shall be as follows:

(a) Upon an appeal from a judgment in a proceeding for the termination of parental rights, the Court of Appeals shall try the cause anew upon the record; and

(b) Upon an appeal in an equitable action or proceeding other than an appeal from a judgment in a proceeding for the termination of parental rights, the Court of Appeals, acting in its sole discretion, may try the cause anew upon the record or make one or more factual findings anew upon the record.

(4) When the Court of Appeals has tried a cause anew upon the record or has made one or more factual findings anew upon the record, the Supreme Court may limit its review of the decision of the Court of Appeals to questions of law.).

Id.

involving the termination of parental rights; this carve-out was rooted in opposition from attorneys practicing family law who believed that such cases were significant enough that they merited plenary fact review by the appellate courts.¹⁴⁶

The Oregon Court of Appeals adopted rules governing the circumstances in which the Court of Appeals will exercise this discretionary *de novo* review process. Pursuant to Oregon Rule of Appellate Procedure 5.40, “[i]n those proceedings in which the Court of Appeals has discretion to try the cause anew on the record and the appellant seeks to have the court exercise that discretion, the appellant shall concisely state the reasons why the court should do so.”¹⁴⁷ When appellants ask the Court of Appeals to change factual findings that a circuit court made below, the appellant is required to “identify with particularity the factual findings that the appellant seeks to have the court find anew on the record and shall concisely state the reasons why the court should do so.”¹⁴⁸ The Court of Appeals has stated that it will “exercise its discretion to try the cause anew on the record or to make one or more factual findings anew on the record only in exceptional cases,” and that requests to do so are “disfavored.”¹⁴⁹ The Court has also set forth several factors it will use in evaluating whether it should exercise its *de novo* authority, including questions about the nature of the trial court’s factual findings and the degree to which the trial court’s findings were incorrect.¹⁵⁰

146. See Holland & Sokol, *supra* note 68, at 22 (noting opposition of family law bar).

147. OR. R. APP. P. 5.40(8)(a).

148. OR. R. APP. P. 5.40(8)(b).

149. OR. R. APP. P. 5.40(8)(c).

150. OR. R. APP. P. 5.40(8)(d) provides:

(d) The Court of Appeals considers the items set out below to be relevant to the decision whether to exercise its discretion to try the cause anew on the record or make one or more factual findings anew on the record. These considerations, which are neither exclusive nor binding, are published to inform and assist the bar and the public.

(i) Whether the trial court made express factual findings, including demeanor-based credibility findings.

(ii) Whether the trial court’s decision comports with its express factual findings or with uncontroverted evidence in the record.

(iii) Whether the trial court was specifically alerted to a disputed factual matter and the importance of that disputed factual matter to the trial court’s ultimate disposition of the case or to the assignment(s) of error raised on appeal.

(iv) Whether the factual finding(s) that the appellant requests the court find anew is important to the trial court’s ruling that is at issue on appeal (i.e., whether an appellate determination of the facts in appellant’s favor would likely provide a basis for reversing or modifying the trial court’s ruling).

(v) Whether the trial court made an erroneous legal ruling, reversal or modification

This summary of the factors emphasizes one point that the appellate courts stressed in requesting this change: the assertion that the Court of Appeals would not intentionally change its rate of reversal in cases that fit within the scope of the statute. According to the chief judge of the Court of Appeals at the time, if the Court had believed that the case merited reversal under the old, *de novo*-required statute, the Court would also choose, under the amended statute, to use its *de novo* review authority.¹⁵¹ Curiously, this meant that, after the amendment, one would generally *not* expect to see cases in which the Court reversed based on a conclusion that there was “no evidence” to support the facts found by the trial court. In those cases, the Court would instead choose to utilize *de novo* review and reverse using the less deferential standard.

Therefore, while the Court of Appeals did not think that its reversal rate would change at all, it did hope that the flexibility provided by the statute would limit the temptation for parties to set out extended factual discussions in their briefs, and that it would also permit the Court to limit its own discussion of the facts in its opinions. The goal, then, was not to change reversal rates, but to improve the efficiency of the Court’s briefing, consideration, and opinion-drafting process.

To test the Court’s predictions regarding the unchanged reversal rates—as well as its predictions regarding the length of factual discussions in briefs and appellate opinions—this Article examines all Court of Appeals cases from 2006 to 2008, and from 2012 to 2013,¹⁵² that cited ORS 19.415, or that referenced the option of *de novo* review in equitable proceedings. For each case, and for the purposes of this Article, it was determined which standard of review the Court used in evaluating the facts of the case, the length of the opening brief’s statement of facts, and the length of any statement of facts in the Court’s opinion.¹⁵³

of which would substantially alter the admissible contents of the record (e.g., a ruling on the admissibility of evidence), and determination of factual issues on the altered record in the Court of Appeals, rather than remand to the trial court for reconsideration, would be judicially efficient.

151. Brewer, CLE Remarks, *supra* note 144.

152. I initially planned on using three years of data on both sides of the 2009 amendments. The amendments only applied to cases filed after the 2009 effective date of the statute, and to avoid any possible effects associated with the transition from one standard to another, I avoided the years immediately surrounding that transition. On further review, however, most of the reported cases decided in 2011 involved appeals filed before the effective date of the amendment, so I excluded cases decided in 2011, and I was unwilling to add 2014 cases because the court expanded by three judges in early 2014, and I believed that the additional judges—not so much the fact that there were different judges, but rather the fact that the court’s capacity expanded—would have complicated the results of the analysis.

153. This necessarily required some judgment in cases where the arguments in the brief, or the

One initial consideration immediately jumps out from the data: nearly twice as many cases reference ORS 19.415(3) in the post-amendment period than in the pre-amendment period.¹⁵⁴ While the Court's caseload did increase to some degree over this period, the increase was not proportional. But there are some potential explanations for this increase. First, the fact of the amendment—and the surrounding discussion of its relevance—as well as the presence of the Court's new rules regarding the discretionary standard have drawn the attention of parties to the potential for *de novo* review. There might be cases in the pre-amendment period in which (at least in theory) *de novo* review might have been appropriate, but in which neither the parties nor the Court chose to raise the question. This Article postulates that all of the “missing” pre-2009 cases—if there are any—are likely to be affirmances. If the Court were to reverse on the alternative “clear error” or “any evidence” standard that would have been applied in the absence of *de novo* review, the parties or court would likely have taken the additional step to realize that *de novo* review would apply under ORS 19.415(3).

One further explanation for at least part of the increase might be found in a change in the nature of cases: in the post-2008 decisions, there were thirty-two juvenile dependency cases listing the Oregon Department of Health and Human Services (“DHS”) as a party and no such cases pre-2008. These post-2008 DHS dependency cases, however, constituted only about half of the total increase in cases referencing ORS 19.415, and their exclusion from the analysis did not alter the outcome below.¹⁵⁵

The results of this analysis demonstrate that the Court of Appeals' prediction—that reversal rates would not change—appears to be correct. This is not a surprise, given that the statutory change was made with the intent of matching statutory language to the Court's existing approach—which already took into account the judges' view of the relative competencies of the courts regarding fact finding. The Court was incorrect, however, in thinking that the amount of time spent on the issue in the parties' briefs would be reduced; in fact, because the Court now asks parties to explain why they believe that *de novo* review is appropriate, parties are effectively writing an entire new section in their

Court's own discussion of the issues on appeal, heavily implicated facts that were discussed for the first time in that argument or ruling.

154. There were thirty-six cases that cited ORS 19.415(3) in the pre-amendment period and seventy-nine cases that referenced it in the post-amendment period.

155. Lora Keenan, a Staff Attorney with the Oregon Court of Appeals, suggested that the change in case makeup was tied to additional funding received in 2009 by the Oregon Protective Services Division to increase the department's ability to litigate appeals in such cases. Notes of Telephone Conversation between Jeffrey Dobbins and Lora Keenan (Nov. 18, 2014) (on file with author).

opening briefs and spending a nearly identical amount of time on their facts. There is no real improvement in the amount of time parties are spending on facts in their briefs—including the time spent on the standard of review—and it appears that the parties are spending *more* time now on facts and the standard of review than was true before the change. The length of the Court’s opinions also demonstrates that the Court—perhaps because it has to spend time explaining why it is choosing a particular standard or because there is more focus on the factual questions—also has to spend more time on the facts of each case.

In the end, the absence of a real signal in the rate of reversal is not particularly surprising. Given the intent of the statute, it might be expected that the Court of Appeals would calibrate its post-amendment rulings to maintain a rate of reversal similar to that prior to the amendment. The point, however, is this: the reversal rate was governed more by the Court of Appeals’ preexisting understandings about the proper responsibilities of the trial court and the reviewing court in the overall appellate system, than it was governed by the precise language of the standard of review. The standard, in other words, does not matter as much as what the standard represents about the relationship between the Court of Appeals and the trial courts. Absent a more fundamental shift in that relationship, one would not expect—and does not see—a change in the rates of reversal.

B. Other Examples of Changing Standards of Review

Are there other circumstances when changes in a standard of review have been imposed from the outside that might be used to test the proposition that standards of review—in the absence of changes in the reviewing court’s perception about the relationship between the relevant decision makers—will not really lead to changes in the underlying reversal rate? There are at least two examples that can initiate the conversation and thought process about this issue, but a statistical examination of these cases will be left for a later Article.

1. Review of Interpretations of Law by Federal Administrative Agencies, 1987–1997, 2002–2012

Students and practitioners of administrative law are familiar with the basic *Chevron* test,¹⁵⁶ which governs the level of deference that federal

156. Arising, of course, out of the United States Supreme Court’s decision in *Chevron*, 467 U.S. 837.

courts give to certain agency interpretations of law.¹⁵⁷ Courts examining the validity of an agency's interpretation of law engage in the so-called "Chevron two-step" deference.¹⁵⁸ Under this algorithm, courts first ask whether Congress has "directly spoken to the precise question at issue" in the case.¹⁵⁹ If Congress has done so, then the statute must be followed, regardless of an agency's contrary view.¹⁶⁰ If the statute is not clear on the relevant point, the second step calls on courts to defer to an agency's interpretation of law as long as it is a "reasonable" interpretation of the underlying statutes.¹⁶¹

Commentators have spilled much ink on the interesting and complicated subtleties associated with *Chevron* deference. Despite those subtleties, however, *Chevron* rapidly became a highly cited United States Supreme Court case. In the first fifteen years after the decision (from 1985 to 1999), the federal Courts of Appeals cited this Supreme Court decision over 2,500 times.¹⁶² Whether because of its straightforward algorithm for determining the appropriate deference to agency legal decisions or for some other reason, Courts took to the *Chevron* algorithm with vigor.

In 2000 and 2001, however, the Supreme Court added a twist to the established *Chevron* algorithm. In *Christensen v. Harris County*, the Supreme Court suggested that not *all* agency interpretations of law were entitled to deference.¹⁶³ In *Christensen*, the Court refused to give deference to a letter opinion issued by the Department of Labor regarding the proper interpretation of the Fair Labor Standards Act, despite ambiguity in the statute and an apparently reasonable interpretation concluding that the informal nature of the letter did not justify *Chevron* deference.¹⁶⁴

Just a year later, in *United States v. Mead Corporation*, the Court formalized the approach that it presaged in *Christensen*, adding a "step

157. *Chevron, U.S.A., Inc. v. Nat. Res. Def. Council, Inc.*, 467 U.S. 837, 845 (1984); see CHARLES H. KOCH & RICHARD MURPHY, ADMINISTRATIVE LAW AND PRACTICE § 11:32 (4th ed. 2016) (discussing judicial review of administrative interpretations of law).

158. See *Peter Pan Bus Lines, Inc. v. Fed. Motor Carrier Safety Admin.*, 471 F.3d 1350, 1353 (D.C. Cir. 2006) (applying "*Chevron* two-step" deference to a decision by the Federal Motor Carrier Safety Administration).

159. *Chevron*, 467 U.S. at 842.

160. *Id.* at 842–43.

161. *Id.* at 843–44.

162. WestlawNext search for citations to *Chevron* in federal Courts of Appeal cases resulted in 2,534 results.

163. 529 U.S. 576, 587 (2000).

164. *Id.*

zero” to the *Chevron* analysis.¹⁶⁵ After *Mead*, reviewing courts are called upon to examine whether an agency has been delegated the authority to speak with the “force of law,” and whether the agency was exercising that authority in making the decision on review.¹⁶⁶ If it is not speaking with the “force of law,” an agency’s decision is entitled only to *Skidmore v. Swift & Company* deference—under which courts defer to the persuasive value of an agency’s determination, but they do not feel themselves bound by it.¹⁶⁷

While the Court majority in *Mead* described the decision as an unremarkable gloss on long-accepted principles associated with *Chevron*,¹⁶⁸ the dissent’s view strongly suggested that the majority opinion marked a sea change in the standard of review for important (though not “speaking with the force of law”) agency decisions.¹⁶⁹ Subsequent cases bear out the dissent’s belief that the decisions in *Christensen* and *Mead* marked a significant change in the Supreme Court’s *Chevron* doctrine, and demonstrate a marked decrease in the Court’s reliance on *Chevron* deference.¹⁷⁰

Before *Christensen* and *Mead*, courts enthusiastically applied *Chevron* deference to agency legal interpretations (almost) regardless of the source.¹⁷¹ After the decisions, courts were called upon to evaluate whether a particular decision was within the scope of statements carrying the “force of law” as delegated to it by Congress.

While the Supreme Court’s decisions in *Mead* and *Christensen* did not

165. See 533 U.S. 218, 237–38 (2001) (holding that United States Custom’s ruling letters were not entitled to *Chevron* deference because Congress did not intend these letters to have the “force of law”).

166. *Id.*

167. *Skidmore v. Swift & Co.*, 323 U.S. 134, 140 (1944).

168. *Mead*, 533 U.S. at 237.

169. *Id.* at 241 (Scalia, J., dissenting).

170. See Linda Jellum, *Chevron's Demise: A Survey of Chevron from Infancy to Senescence*, 59 ADMIN. L. REV. 725, 781 (2007) (noting how “*Chevron*’s stronghold appears to be weakening”).

171. Not all federal courts did so; the federal circuit’s refusal to defer to the Customs Service’s tariff rulings was the source of a circuit split that was subsequently resolved by *Mead*. As the United States Solicitor General noted in the petition for certiorari in *Mead*, most federal courts of appeals at that time held that “agency interpretations adopted by means other than formal regulations are entitled to judicial deference.” Petition for Writ of Certiorari at 7–18, *Mead*, 533 U.S. 218 (No. 99-1434) (citing, e.g., *Ass’n of Bituminous Contractors, Inc. v. Apfel*, 156 F.3d 1246, 1251–52 (D.C. Cir. 1998)); *Gould v. Shalala*, 30 F.3d 714, 719–20, 720 n.7 (6th Cir. 1994); *Georgia Dep’t of Med. Assistance v. Shalala*, 8 F.3d 1565, 1571 n.8 (11th Cir. 1993); *Coca Cola Co. v. Atchison, Topeka, & Santa Fe Ry.*, 608 F.2d 213, 215 (5th Cir. 1979); also citing *Elizabeth Blackwell Health Ctr. for Women v. Knoll*, 61 F.3d 170, 181-182 (3d Cir. 1995) (according deference to agency policy statement), *cert. denied*, 516 U.S. 1093 (1996); *Emerson v. Steffen*, 959 F.2d 119, 122 (8th Cir. 1992) (same)).

purport to be anything other than a slight modification of the *Chevron* approach, the opinions certainly altered the approach with which the lower courts had become familiar. With few exceptions, most lower courts had not previously distinguished between the deference to be applied to agency statements that did not “speak with the force of law,” and statements that did so; the distinction simply was not part of the analysis. Once the Supreme Court introduced “step zero,”¹⁷² however, courts were called upon to examine whether an agency’s decision spoke with “the force of law,” and if not, *Skidmore* deference, rather than *Chevron*, would be applied.¹⁷³

An examination into the rates of reversal in decisions reviewing agency actions speaking *without* the “force of law” should reveal the degree to which this change in the standard of review for agency decisions in this category altered the rates of reversal in the United States Courts of Appeal.

The analysis would require examining United States Courts of Appeal opinions from the pre-*Christensen* period to identify circumstances in which agency statements of law are articulated in “subzero” cases (i.e., cases where the analysis would not move beyond “step zero” under the law as stated after the 2002 *Mead* decision.¹⁷⁴ To avoid possible confounding effects that the petition (and certainly its grant) might have on the decisions of the relevant Circuit Court of Appeals, the analysis would have to focus on a United States Court of Appeals that adhered to a “pure” *Chevron* analysis for all agency decisions in the period predating the filing of the petition for certiorari in *Christensen*.¹⁷⁵ The pre-*Christensen* reversal rate could then be compared against the post-*Mead* reversal rate in those same appellate courts over a similar time period.¹⁷⁶

172. This is the shorthand used by commentators and courts to refer to the step introduced by *Christensen* and *Mead*. See Cass R. Sunstein, *Chevron Step Zero*, 92 VA. L. REV. 187, 190–92 (2006).

173. *Id.* at 211–12; see also *Prime Time Int’l Co. v. Vilsack*, 930 F. Supp. 2d 240, 248 (D.D.C. 2013), *aff’d sub nom.* *Prime Time Int’l Co. v. U.S. Dep’t of Agric.*, 753 F.3d 1339 (D.C. Cir. 2014) (“The threshold inquiry—sometimes called *Chevron* ‘step zero’—is determining whether Congress has delegated interpretive authority to the agency in question.”).

174. It might also be informative, though certainly more work, to categorize all relevant cases decided using *Chevron*’s algorithm, even if they would not fit into the “subzero” categories.

175. The Fifth Circuit decision that was under review in *Christensen* was issued in October, 1998; so, an ideal span of decisions for review would involve decisions issued between July 1, 1996 and July 1, 1998. Determining the best circuits for this purpose will take additional investigation into how they approached this question in the years preceding *Christensen*; the Federal Circuit is the Court of Appeals that issued the (non-deferential) decision that later came under review in *Mead*, so that court is not a candidate for the study.

176. To match the period from before *Christensen*, I would use July 1, 2003 (to allow news of

Again, although this would not eliminate all confounding factors as judges and cases change, presumably the makeup of the relevant court and its case load would be similar in the two time periods, which might eliminate some confounding factors that could otherwise influence reversal rates from very different cases or courts.

2. Review of Downward Departures Under the Federal Sentencing Guidelines, 1991–1997

A second change that might merit future investigation is the level of discretion that federal courts use to review decisions by lower courts that depart from baseline sentences calculated under the Federal Sentencing Guidelines. Although subsequent United States Supreme Court decisions in *Blakely v. Washington*¹⁷⁷ and *United States v. Booker*¹⁷⁸ have since undermined the binding nature of the Federal Sentencing Guidelines, the period between their initial promulgation and the “beginning of the end” of their mandatory application in federal court is one in which the standard of review changed on several occasions. For purposes of this analysis, this Part of the Article focuses on the shift occasioned by the United States Supreme Court’s decision in *Koon v. United States*, which effectively eased judicial review of departures in most circuits.¹⁷⁹

At their heart, the Federal Sentencing Guidelines established an algorithm for determining the appropriate sentence for a particular defendant based on the nature and circumstances of his offense and a variety of other factors.¹⁸⁰ In rare cases, district court judges were given the ability to “depart” from the sentence mandated by the Federal Sentencing Guidelines.¹⁸¹ District courts only rarely exercised this authority, but when they did, appellate courts were tasked with reviewing the appropriateness of that departure. Until the decision in *Koon*, the federal appellate courts generally reviewed the validity of departures via a staged approach under which any facts justifying the departure were examined using a “clear error” standard, while the decision whether to

Mead to “sink in” to the minds of practitioners and judges) through July 1, 2006.

177. 542 U.S. 296, 304–05 (2004) (holding that Washington State’s sentencing statute violated the Sixth Amendment’s jury trial guarantee).

178. 543 U.S. 220, 244 (2005) (holding that, except for prior convictions, all facts used to apply federal sentencing guidelines must be proven to a jury beyond a reasonable doubt).

179. 518 U.S. 81 (1996); Fischman & Schanzenbach, *supra* note 94, at 412 (referring to *United States v. Rivera*, 994 F.2d 942 (1993)).

180. See 18 U.S.C. § 3553(a) (2012) (citing baseline sentencing principles, including particular characteristics of the offense and the defendant).

181. See *id.* § 3553 (noting instances where judges may use discretion within the mandated Federal Sentencing Guidelines).

depart was reviewed de novo, and the magnitude of the departure was reviewed for abuse of discretion.¹⁸² Although the First Circuit was the first to adopt this tiered approach to departure decisions, it was also the first to abandon it, thereby setting up a circuit split that led to the Supreme Court's decision in *Koon*.¹⁸³

In the 1996 decision in *Koon*, the United States Supreme Court adopted a reasonableness standard for the entire departure decision—calling on the United States Courts of Appeal to apply an “abuse of discretion” test to determine whether departures were valid.¹⁸⁴

Some work has already been conducted into the effect of this change (as well as subsequent changes) on the rate of reversal in the United States Courts of Appeal. For example, in 2002, Paul Verkuil's study briefly examined sentencing decisions, but he did so only for the post-*Koon* era in an effort to evaluate the reversal rate under the deferential “abuse of discretion” standard.¹⁸⁵ He concluded that the reversal rate was approximately 20 percent, but did not compare that rate to the pre-*Koon* era during which the tiered review process applied de novo review to the decision to depart,¹⁸⁶ so it provides little information about whether the change in the standard of review resulted in a shift in outcomes. Some authors predicted that it would not.¹⁸⁷

As noted above, Fischman and Schanzenbach concluded in 2011 that district court judges were affected by the change in the standard of review for departures under the guidelines, and they assumed that this meant that appellate courts were also treating cases differently.¹⁸⁸ But this conclusion was not directly measured; it was simply presumed from the

182. *United States v. Rivera*, 994 F.2d 942, 951–52 (1st Cir. 1993); Fischman & Schanzenbach, *supra* note 94, at 412.

183. *United States v. Diaz-Villafane*, 874 F.2d 43, 51–52 (1st Cir. 1989).

184. *Koon*, 518 U.S. at 98–100. This standard was subsequently changed by statute, and again by the Supreme Court's decision in *United States v. Booker*, 543 U.S. 220 (2005), which effectively removed all mandatory features of the federal sentencing guidelines to guarantee compliance with the Sixth Amendment. *See* Prosecutorial Remedies and Other Tools to End the Exploitation of Children Today Act of 2003, Pub. L. No. 108-21, § 401, 117 Stat. 650, 667 (codified as amended in scattered sections of 18 U.S.C.).

185. Verkuil, *supra* note 16, at 720–22.

186. *Id.*

187. *See* Fischman & Schanzenbach, *supra* note 94, at 406 (noting the scholars that doubt that “changing from a de novo to an abuse of discretion standards would have much impact”). To prove their point, their article cites generally to Cynthia K. Y. Lee, *A New “Sliding Scale of Difference” Approach to Abuse of Discretion: Appellate Review of District Court Departures under the Federal Sentencing Guidelines*, 35 AM. CRIM. L. REV. 1 (1997) and KATE STITH & JOSE A. CABRANES, *FEAR OF JUDGING: SENTENCING GUIDELINES IN THE FEDERAL COURTS* (1998). *Id.*

188. *See supra* text accompanying notes 94–99 (describing the 2011 study by Fischman and Schanzenbach).

fact that district court decisions were changing.¹⁸⁹ A more careful study of appellate cases under other circumstances of changed standards of review should provide other opportunities to fully evaluate the effect of the standard of review on the work of appellate courts.

III. LESSONS FROM CHANGE AND SUGGESTIONS FOR CHANGE

Standards of review are, at their heart, statements about the proper relationship between structural components of the appellate process. Discussions about the appropriate standard of review are one of the few places that our appellate system implicitly (if not explicitly) discusses this relationship. If one compares the discussion about standards of review, for instance, to the general lack of discussion in appellate cases about the bindingness of precedent,¹⁹⁰ a standard of review discussion includes much more analysis into the roles of appellate courts, trial courts, and agencies, as well as into the relative competence of those entities in various parts of the legal process.

This Article suggests, then, that the value of discussions about standards of review comes not so much from a direct effect of the standard on the outcome in a particular case, but from the rhetorical effect that the standard—and the implications for appellate structure that are embodied within it—has on how the reviewing courts (as well as the advocates, lower courts, and agencies) perceive their role. While there does appear to be a net effect from the application of a particular standard in a particular case, this Article suggests that this results not so much from applying a particular standard to a particular case, but rather to the way in which standards of review affect (and reflect) the way that courts, agencies, and advocates think about their place in the system of legal review.

This perspective may seem rather theoretical, and arguably irrelevant: If standards of review matter—and they do, at least rhetorically—then why does it matter *how* they matter? An appellate court might reverse a trial court's evidentiary ruling by explicitly concluding that it amounted to an abuse of discretion. Alternatively, it might reverse because it *calls* something an abuse of discretion because it believes that, given its role in the legal system under the circumstances of that particular case, it has the ability to place more weight on *its* understanding of the proper ruling, rather than the trial court's decision that was made on-the-fly at trial.

189. See *supra* text accompanying notes 94–98 (noting that the 2011 study by Fischmann and Schanzenbach was based on a perception). Again, a direct examination into the data surrounding this change awaits further development.

190. See generally Dobbins, *Structure*, *supra* note 11 (recognizing the difficulty in ascertaining the precedential and binding effect when changes occur within federal appellate review).

Does the difference in the underlying rationale matter?

It appears, however, that the implications for appellate practice and how one thinks about standards of review are tangible and twofold. First, when legislative entities (or rules committees or courts) define standards of review for a particular question, they should recognize that the value of the standard does not come primarily from the definition itself. The precise phrasing of the standard might not make a critical difference to the outcome of a particular case. What is more important is the surrounding discussion (or understanding) regarding the relative role that the reviewing court has vis-à-vis the work of the trial court or the administrative agency. For that reason, those defining the scope of a standard of review should make an effort to incorporate into their discussion (or the legislative or administrative history) an explanation for why it is that the relative authority of the reviewing court should have the relationship to the decision-making body that it does. Whenever standards of review change (or are initially established), this consideration should accompany the development and articulation of that standard.

Second, it suggests that counsel should not be sanguine about standards of review, even in cases where existing law appears to clearly define the relevant standard. As the uncertain data regarding the effect of standards on reversal rates suggests, appellate courts that rely on settled standards of review still have a very broad decision space within which they can apply that standard—a decision space which may well encompass both affirmance and reversal. The responsibility (and opportunity) for counsel in these cases is to communicate, in some way, the underlying rationale and structural justification for the relevant standard in a way that helps to direct the appellate court to an understanding of not so much the standard itself, but the way in which the appellate court thinks about its relative competence compared to that of the decision maker under review.

To come full circle, then, this Article suggests that standards of review *do* matter—but not how one might think. Rather than *directing* a particular outcome in a particular case, the power of the standard of review comes from the rhetorical signal that they send about the relationship and relative competence of the players in the larger legal system. The most useful application of that standard comes from what it suggests about the relative authority of the reviewing court when compared to the deciding court. With that understanding, legislative bodies, courts, and counsel could refocus their attention on standards of review or change them in a manner that fully reflects their underlying purpose. Further examination into other circumstances in which

standards of review have changed at the instigation of judicial decisions (like those in *Koon* and *Mead*) should shed further light on the lessons that can be drawn from changing standards of review.

TABLE 1¹⁹¹

		Outcome on Questions of Fact		Length of Case Facts
		Affirmed	Reversed	(words)
Pre-Amendment (all de novo review)	N=36	71%	29%	1052
Post-Amendment (discretionary de novo review)	Overall N=79	75%	25%	1438
	“Any Evidence” N=66	78%	22%	
	De Novo* N=13	61%	28%	

*Of the post-amendment de novo review cases, six (6) were termination of parental rights cases, in which de novo review is still mandated, all of those were affirmed.

191. There is no significant difference in reversal rates overall, between pre-amendment and post-amendment using “any evidence,” and between pre-amendment and post-amendment using “de novo.” There is a significant difference ($p < .05$) in the length of the fact discussion with the length greater in post-amendment decisions.