Norms, Empiricism, and Canons in Statutory Interpretation

William N. Eskridge, Jr.†

In the abstract, there should be little quarrel with the proposition that theories of statutory interpretation should be subjected to empirical testing.[1] Pragmatic thinkers and practitioners ought to be open to any theory that has a "cash-value," as philosopher William James put it.[2] If an uncompromising application of statutory plain meaning, what I have called the "new textualism," produces more predictable results than eclectic theories now deployed by pragmatic judges and scholars, then theory and practice ought to move toward textualism, even more than they already have.[3] If studies deploying a variety of rigorous techniques show that following a textualist approach generates more consistent and uniform interpretations by a variety of different judges than other approaches to statutory interpretation (original intent, purpose), pragmatists ought to consider those findings as a significant reason to support and adopt a more strictly textual approach to statutory interpretation.

Many formalists appreciate the Jamesian insight. Text-based theories of statutory interpretation are usually justified in part by factual claims. For instance, Justice Antonin Scalia argues that judges should not consult legislative history because such materials (1) are an illegitimate source of statutory meaning under Article I, Section 7 of the Constitution, which renders the text the only source of law and augurs against crediting the views of congressional subgroups; (2) are unreliable and often misleading evidence of legislative deals, and hence are not worth the re-

[1] For examples, see Cass R. Sunstein, *Must Formalism Be Defended Empirically?*, 66 U Chi L Rev 636 (1999); William N. Eskridge, Jr., *Textualism, The Unknown Ideal?*, 96 Mich L Rev 1509, 1541-42 (1998).

[2] "Pragmatism . . . asks its usual question. 'Grant an idea or belief to be true,' it says, 'what concrete difference will its being true make in any one's actual life? . . . What, in short, is the truth's cash-value in experiential terms?'" William James, *Pragmatism: A New Name for Some Old Ways of Thinking* (Longmans 1907), excerpted in John J. McDermott, ed, *The Writings of William James* 430 (Chicago 1977).

[3] William N. Eskridge, Jr., *The New Textualism*, 37 UCLA L Rev 621 (1990).

search expended by lawyers and judges; and (3) increase the discretion of willful judges to read statutes to reflect their own preferences.[4] Only the first is a purely formalist argument for Scalia's formalist theory; the latter two mix functional and normative arguments for the theory, and the functional aspect of each might be tested empirically or quasi-empirically. Like Scalia, most of the other leading statutory formalists have provided functional as well as formal reasons in favor of their theories.[5]

Subjecting theories of statutory interpretation to systematic factual testing, therefore, should be potentially attractive to both formalists and pragmatists. Empirical work might expand arenas of consensus among scholars and judges about the latter's appropriate role and reasoning in statutory cases. For example, if testing revealed that purpose-based theories of statutory interpretation yielded completely unpredictable applications by different judges, support for purposivism as a general theory would diminish. Even if such work did not immediately eliminate arenas of disagreement, the results would be intrinsically interesting and illuminating—we would know more about the application of statutes by interpreters. Additionally, empiricism might improve the quality of the normative debate, if it grounded arguments in concrete reality and forced each side to understand the weaknesses or gaps in its respective arguments. Finally, descriptive work that does not contribute to current normative debate is potentially useful. As we have learned from the natural sciences, data irrelevant now might become relevant to future debates and might even impel scholars to reconsider the focus of debate.

On the other hand, it is not clear that empirical testing will eliminate or even reduce normative disagreement. It might even

---

[4] Antonin Scalia, *A Matter of Interpretation: Federal Courts and the Law* 29-37 (Princeton 1997). See also John Manning, *Textualism as a Nondelegation Doctrine*, 97 Colum L Rev 673 (1997) (arguing that judicial reliance on legislative history allows bill sponsors and committee members in effect to say what a law means, allowing other members of Congress to express dissenting understandings of the law, thereby escaping the negative political consequences of having voted for it).

[5] See, for example, Adrian Vermeule, *Legislative History and the Limits of Judicial Competence: The Untold Story of* Holy Trinity Church, 50 Stan L Rev 1833, 1891-95 (1998); Martin H. Redish and Theodore T. Chung, *Democratic Theory and the Legislative Process: Mourning the Death of Originalism in Statutory Interpretation*, 68 Tulane L Rev 803, 871-73 (1994); Lawrence C. Marshall, *"Let Congress Do It": The Case for an Absolute Rule of Statutory Stare Decisis*, 88 Mich L Rev 177 (1989); Earl M. Maltz, *Statutory Interpretation and Legislative Power: The Case for a Modified Intentionalist Approach*, 63 Tulane L Rev 1 (1988); Frank H. Easterbrook, *The Supreme Court, 1983 Term: Foreword: The Court and the Economic System*, 98 Harv L Rev 4, 14-18, 42-59 (1984).

increase such disagreement. Nor is it inevitable that empiricism will increase our stock of factual information. The reason is that facts and values cannot always be disentangled. Consider the phenomenon at a general level. Presumptions are critical in normative argumentation: Where you start is usually where you end up. Presumptions affect or even determine the cash-value of factual findings. A theorist for whom textualism is presumptively the best normative theory (Justice Scalia) will be more easily satisfied with equivocal empirical support for the theory than would a theorist for whom other theories are presumptively better (Justice Breyer). Empirical studies are typically beset with questions about whether their conclusions are reliable, significant, and generalizable. When there is room for quarrel or dubiety, presumptions will either trump empirical findings or, more subtly, influence the way the reader understands as well as evaluates such findings. If you are strongly committed to your normative starting point, that commitment will not only raise the bar for contrary evidence to change your mind but will pervasively influence the way you read that evidence.

The nature of presumptions also renders the factual inquiry potentially regressive, and either circular or endless. Presumptions themselves are a mix of factual and normative considerations. A statutory interpreter who thinks it morally important to find and apply the original expectations of the enacting legislature is also more likely to believe that such an imaginative reconstruction is actually possible most if not all the time. Conversely, an interpreter who believes that imaginative reconstruction is possible or even easy in the typical case is more likely to find the enterprise valuable. The factual features of a presumption might be tested empirically, but can they be separated from the normative features? Not immediately, opined William James. Because our mind approaches matters holistically rather than compartmentally, such separation is not usually possible, although James believed minds can change slowly as they adapt to new information.[6]

These points about presumptions are not only generally applicable to statutory interpretation discourse, but specifically relate to the main body of statutory interpretation doctrine, namely, the canons of statutory construction. The canons are on their face formal presumptions or rules about statutory meaning.

---

[6] James, *Pragmatism* at 382-83, 418-19 (cited in note 2).

In our collection of Supreme Court canons, Philip Frickey and I suggested that canons belong to three different categories.[7] Some are *linguistic* presumptions about what statutory language means. For example, the canon *inclusio unius est exclusio alterius* (the inclusion of one thing suggests the exclusion of all others) presumes that, when Congress composes lists of matters included in a statute, omitted matters are not covered by the statute. Other canons are presumptions about *extrinsic sources*. Canons about deployment of legislative history fall into this category, but the most important extrinsic source canon is the presumption or rule that courts will defer to agency interpretations of statutes they are charged with enforcing, unless Congress has clearly legislated otherwise. Most canons are neither linguistic nor extrinsic source rules but are *substantive* in nature. The best example is the rule of lenity, which presumes that ambiguous penal statutes will be construed against the state.

The force of the canons depends upon the willingness of judges to internalize their presumptions and not just deploy the canons as window-dressing for interpretations reached on other grounds. There has long been a vigorous debate over the utility of the canons. They have been reviled by the leading statutory theorists,[8] but they dominate judicial opinions and have found a new generation of fans.[9] After decades of discussion and experience, it is far from clear that the debate has been resolved or even been advanced. Would empirical testing advance or resolve some of the debate? Consider three normative defenses of the canons, the empirical issues suggested within them, and the hidden normative commitments that might impede consensus even if the data prove one-sided.

---

[7] William N. Eskridge, Jr. and Philip P. Frickey, *The Supreme Court, 1993 Term: Foreword: Law as Equilibrium*, 108 Harv L Rev 26, 65-71, 97-108 (1994); William N. Eskridge, Jr. and Philip P. Frickey, *Quasi-Constitutional Law: Clear Statement Rules as Constitutional Lawmaking*, 45 Vand L Rev 593 (1992).

[8] Notable exceptions include Henry M. Hart, Jr. and Albert M. Sacks, *The Legal Process: Basic Problems in the Making and Application of Law* 1191 (Foundation 1994) (William N. Eskridge, Jr. and Philip P. Frickey, eds); Samuel Mermin, *Law and the Legal System: An Introduction* 264-65 (Little, Brown 2d ed 1982).

[9] See, for example, William N. Eskridge, Jr., *Dynamic Statutory Interpretation* 275-306 (Harvard 1994); Geoffrey P. Miller, *Pragmatics and the Maxims of Interpretation*, 1990 Wis L Rev 1179; Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 Harv L Rev 405, 451-62 (1989).

1. Democracy Justifications

A most attractive justification would be that the canons supply details to statutory schemes that the legislature would probably have provided if its members had written a more detailed statute.[10] In our polity, statutory legitimacy is closely linked to representative democracy. The people participate in lawmaking indirectly but powerfully by choosing representatives who reflect their preferences and by monitoring the representatives' performance through subsequent elections. If courts follow rules of interpretation that faithfully reflect the preferences of those representatives (and, indirectly, of the citizenry as well), those rules have a strong claim to legitimacy on the basis of the democratic values they serve.

If I have put the issue properly, this would appear to be a case where empirical testing would be useful and might narrow or even resolve disagreement. There is overlapping consensus about the central normative claim that statutes should have a close tie to representative democracy: the claim finds support in the Constitution as well as contemporary popular attitudes, in America's historical experience as well as political theory, among liberals and conservatives alike. To evaluate the canons under this uncontroversial criterion requires factual judgments, and academic detective work might be able to inform or even establish these factual judgments.

Yet even in cases like this, where there is much normative agreement, it will typically be hard to devise a proper empirical test. For example, if legislators should find congenial the canon of deference to agency interpretations as a general matter, this canon might be defensible on democracy grounds. But how could that proposition be tested? One could poll members of Congress, but cooperation might not be forthcoming, and a current poll would not tell us what prior Congresses thought. It would also be hard to pose the issue reliably. A member might say yes to the inquiry, "Should courts generally defer to agencies?," while also saying yes to the inquiry, "Should courts refuse to defer to unreliable agencies or agencies that are ideologically slanted?" While

---

[10] Debate over this justification is usually set at a general level: Congress generally, and not the specific Congress in question, would fill in statutory *lacunae* in the specified way. See Scalia, *A Matter of Interpretation* at 28-29 (cited in note 4) (critiquing rules of construction); Sunstein, 66 U Chi L Rev at 645-46, 653-55 (cited in note 1) (discussing interpretive rules that "mimic" the market).

perhaps not insuperable, these and other conundrums complicate the project of empirical testing.[11]

Often, the problem with empirical testing will be normative as well as technical. Consider the *inclusio unius* canon. Although this is one of the most frequently invoked linguistic canons, it strikes me as an unreliable rule of thumb about the ordinary use of language. My judgment might be tested, without insuperable design problems. Linguists have gained valuable experience conducting empirical studies to determine how words are regularly used, and they have recently teamed up with law professors to figure out the ordinary meaning of statutory terms and sentences.[12] In everyday parlance, are lists considered exhaustive? My hypothesis would be that *inclusio unius* is only sometimes a reliable maxim, and whether it's reliable depends on normative baselines in the particular case. For example, if Sally is kicking and pushing her little brother and Mother says, "Stop kicking and pushing your brother," Sally has no valid *inclusio unius* argument authorizing her to choke her hapless sibling. The list in the directive would be read to stand for a more general principle (no hurting) that would be fairly enforceable against Sally. On the other hand, if Mother tells Sally, "You can have one cookie and one scoop of chocolate ice cream," *inclusio unius* is applicable if Mother strictly regulates the children's consumption of sweets. The list in the directive would be read to stand for a limited authorization for Sally to depart from the general principle (no sweets).

The foregoing thought experiments suggest a deep problem with testing legal canons and theories. Empirical testing is only worthwhile if set at a general level: "Do ordinary speakers usu-

---

[11] Personally, I would find a casually empirical approach illuminating. Casual empiricism would include interviews with members of Congress and their staffs and accounts by former players in the legislative process. See, for example, James J. Brudney, *Congressional Commentary on Judicial Interpretations of Statutes: Idle Chatter or Telling Response?*, 93 Mich L Rev 1 (1994) (drawing in part on the comments of former members of Congress and on the author's experience as a Senate staffer to argue that courts should look to legislative history in interpreting statutes). Such materials would have to be discounted in terms of representativeness, but they could provide us with thicker descriptions of the process and a much better basis for judgment. To have a cash-value, evidence does not have to dispel all doubt; a story that changes my probability assessment from 35 percent to 40 percent has a significant cash-value.

[12] See, for example, Clark D. Cunningham, et al, *Plain Meaning and Hard Cases*, 103 Yale L J 1561 (1994). Empirical canvassing of how words are regularly used is now showing up in judicial opinions. See, for example, the debate between the majority and the dissent in *Muscarello v United States*, 118 S Ct 1911 (1998).

ally imply exclusion of omitted items when they compile lists?" Or, "Does Congress generally want courts to defer to agencies?" Even if the answer to these questions is generally "yes," it is always going to be a "yes, but"—yes, but not if inconsistent with the overall statutory policy or constitutional values or human reason. The establishment of a "yes, but" presumption has accomplished little or nothing, for it just shifts the inquiry about where the presumption lies from the general to the particular, and the particular cannot be predicted. Even if *inclusio unius* is the correct presumption most of the time, the statutory interpreter ought not rely on it unless she engages in the statutory purpose inquiry and, in some cases, in a more ambitious constitutional inquiry.

Another complexity arises from the institutional context within which statutes are enacted. Because statutes are adopted by legislatures acting under a procedurally and politically messy system of bicameralism, executive presentment, and multiple veto-gates,[13] any testing of the *inclusio unius* maxim should consider the institutional context. When statutes contain lists, we should ask not only, "Would the typical legislator expect no new items to be added in the context of this statutory policy?," but also, "Does the institutional context militate against comprehensive lists?" Law professors could work with political scientists to explore the latter question empirically. But I doubt that even the most rigorous such examination would eliminate, and it might not even reduce, the key role played by normative baselines and the context of the directive. It may be that *inclusio unius* is most appropriately applied to a lengthy list of exclusions from a statute, if the legislature debated the list and edited it through amendment. This proposition is agreeable, if at all, for reasons that are normative (legislative deliberation makes me more willing to stick to the letter of the statute) as much as factual (this is how the typical legislator would understand her text to be applied). And, as before, the particular normative context of the statute being interpreted will make a difference. A list of actors covered by Title VII will likely be read more expansively (civil rights laws should be liberally applied) than a list of actors covered by a criminal statute (rule of lenity).[14]

---

[13] By "vetogate" I mean a place within a process where a statutory proposal can be vetoed or effectively killed.

[14] The statement in the text is itself subject to, and thereby illustrates, the regressive nature of the enterprise. Whether Title VII is read broadly is going to depend on whether the interpreter believes the law has already been too expansively construed, whether the breadth

A final problem is that even if the canons individually or collectively (discussed below) reflect the "typical" legislature's preferences, the canons might not be desirable under democratic theory. Such theory might posit that legislatures depending on courts to fill in the details of statutes are lazy organs and that democracy demands limits on the degree of legislative delegation. (This, in fact, is one idea underlying the rule of lenity for construing criminal statutes.) Thus, further examination is needed to determine whether a comprehensive and preference-reflecting canonical regime saps legislative energy or has countervailing advantages, such as freeing up legislatures to do more productive work. (That in turn might be objected to. Would legislatures do work or just accomplish more mischief?) Is a canonical regime better than its alternatives, such as a chaotic ad hocism? Conversely, the best regime might be one where most of the canons reflected legislative preferences, but a few loose canons did not, the better to bestir legislative deliberation in areas such as the criminal law. (The rule of lenity is probably a loose canon.) And so on. In short, testing begets more testing, and a new cottage industry is created.

2. Rule of Law Justifications

An important value also widely accepted in our polity is the desirability of the rule of law. The rule of law requires that statutes—whatever their source, be it a representative legislature, a plebescite, or a monarch—be applied in an objective, consistent, and transparent way to citizens and others subject to the state's authority. Courts are the guardians of the rule of law, but also a threat to it. An independent judiciary may make it more likely that a statute will be applied in the same way tomorrow as today, and to the powerful and influential as well as to the miserable and the obscure, but an independent judiciary also poses a risk that judges will bend statutes to reflect their own political preferences. The latter would undermine our ability to predict how a statute will be applied, and the cynical among us expect that judicial bending would be slanted in favor of persons or groups the judge identifies with or likes.

The canons of statutory interpretation can be defended if they generate greater objectivity and predictability in statutory

---

seems fair in the particular case, and so forth. The same messiness affects the rule of lenity supposition for the list in a criminal statute.

interpretation. The rule of lenity, for example, might assure the squalid as well as the virtuous that their activities would not land them in jail unless clearly covered by a criminal statute. More important, as a collection of rules the canons might constitute an interpretive regime that both restrains judges and enables the citizenry to predict how those judges will apply ambiguous as well as clear statutes. Not least important, such an interpretive regime could serve democracy values (discussed above) as legislators and their staffs could predict how different proposed statutory language would be applied.[15] Thus, even if ordinary legislative readers do not follow *inclusio unius* or the rule of lenity in processing texts, the judiciary can induce these conventions by putting the legislature on notice that these rules will be followed when courts apply statutes.

Few would object to the overall goal of making the law more predictable, objective, and so forth. The debate has been whether the canons actually constrain judges and, in turn, make statutory application more predictable. Karl Llewellyn's famous thrust-and-parry array maintained that for every canon there is a countercanon; the canons therefore have no independent value in statutory interpretation and are just window-dressing for results reached for other reasons.[16] The large majority of Llewellyn's competing canonical couplets are presumptions about language and extrinsic sources, followed by qualifications to the presumptions. A statute cannot go beyond its text—unless needed to effect the statutory purpose (Llewellyn's first thrust and parry). At one level this is unremarkable. What rule does not have exceptions? At a deeper level, Llewellyn is making the same criticism of the rule of law justification for the canons that I earlier made of the democracy justification: The policy and history of the particular statute and the facts of the particular case will determine or influence whether the rule or the exception is the starting point. Notwithstanding this point, the canons still might be defended on rule of law grounds. Llewellyn ignored the substantive canons, including the rule of lenity. By the thrust-and-parry display of individual canons and countercanons, Llewellyn deflected attention from the most likely value of canons—to work together as an interpretive regime, yielding more consistent results than would exist in its absence.

---

[15] See William N. Eskridge, Jr. and John Ferejohn, *Politics, Interpretation, and the Rule of Law*, in Ian Shapiro, ed, *The Rule of Law* 265 (NOMOS 1994).

[16] Karl N. Llewellyn, *Remarks on the Theory of Appellate Decision and the Rules or Canons About How Statutes Are to be Construed*, 3 Vand L Rev 395, 401-06 (1950).

Perhaps, empirical work can illuminate the debate Llewellyn started. Some questions that might be investigated include the following:

- Would one hundred judges told to follow the canons in an ordinary case be more likely to reach the same result than one hundred judges provided with only more general guidance (such as an instruction to apply statutory plain meaning, legislative intent, or statutory purpose)? Do attorneys find it easier to predict how judges will construe statutes if they understand the canonical interpretive regime?
- Is the legislature or its staff aware of the canons, either individually or as a collective regime of rules and presumptions? If not, would awareness of the canons make statutory drafting easier? Would legislators have better purchase on how their statutes would be applied?
- Is there another interpretive regime (for example, another list of canons or another way of presenting them) that would better constrain judges or allow others to predict their behavior?[17]

These are only the most obvious questions, but they illustrate the above-mentioned problems with the process of empirical testing: How to do the tests? Are all relevant variables controlled for? Can results be generalized, or is the interpretive enterprise too particularistic?

The most useful rule of law test of the canons, and of more general theories of statutory interpretation such as the new textualism, would be: Do they constrain judges or make interpretation more predictable? It would be very hard to develop a good test of this proposition. How do you constitute the sample?[18] What method is required to create a random sample? How can you choose a case that will be both ordinary and representative of many other cases? All of these quandaries could theoretically be

---

[17] This last item would be especially hard to test, because the order in which the canons are considered might affect the results. For example, starting with the rule of lenity might yield somewhat different consequences than ending up with it, perhaps reducing it to tie-breaker status.

[18] Real judges will not sit still for such an exercise, and they are already set in their ways (therefore a corrupted sample). Law students would seem to be less corrupted by preexisting bias, but their very lack of experience would make testing more arduous, for one group would have to be schooled in the canons, another group in plain meaning, and so forth. The schooling process might itself yield corrupting factors.

solved over time and with repeated testing. Who has the re-
sources to do this?[19]

The who-will-do-it quandary is made more acute by hidden
normative issues that threaten the relevance of even the most
rigorous empirical findings. I have treated rule of law values as
uncontroversial in our polity, but there is normative disagree-
ment about how highly to value predictability and even objectiv-
ity in law. Specifically, there is vigorous disagreement as to
whether the law has any room for equity and mercy, whereby the
wise but harder-to-predict judge does justice in the individual
case.[20] In Marguerite Yourcenar's celebrated account, the Roman
Emperor Hadrian opined that the future of the state depended
upon construction of the laws that allowed for mercy. As an exam-
ple, Hadrian cited a case where he sentenced a slave who had at-
tacked him with a knife. Although the attack would have carried
the death penalty had "the law been applied with savage rigor," the
emperor applied the law leniently and thereby gained the loyalty of
an able servant.[21] Many of us admire Hadrian's judgment and, in-
deed, believe that it reinforced rather than undermined the rule of
law. Unfortunately, the goal of the rule of law itself might be in
play: Is the completely predictable judge desirable? At what price
in sacrificing individuated justice?

The same point could be made about democracy. The demo-
cratic value potentially served by the canons needs to be tem-
pered by the observation that our polity might not want statutory
interpretation always to mimic the results reached or that would
have been reached by the legislature. Alexander Hamilton argued
for an independent judiciary as "an essential safeguard against
the effects of occasional ill humors in the society" that lead to the
enactment of "unjust and partial laws."[22] He specifically urged
judges to "mitigat[e] the severity and confin[e] the operation of
such laws."[23] While our polity values judicial deference to demo-

---

[19] In his oral remarks, Cass Sunstein said that Chicago does.

[20] See the classic debate among fictional judges in Lon L. Fuller, *The Case of the Spe-
luncean Explorers*, 62 Harv L Rev 616 (1949). For a wonderful recent example, see the de-
bate between Judges Easterbrook and Posner in *United States v Marshall*, 908 F2d 1312,
1314-26, 1331-38 (7th Cir 1990) (en banc), affd as *Chapman v United States*, 500 US 453
(1991).

[21] Marguerite Yourcenar, *The Memoirs of Hadrian* 116-17 (Farrar, Straus and Young
1954) (Grace Frick, trans).

[22] Federalist 78 (Hamilton), in Clinton Rossiter, ed, *The Federalist Papers* 464, 470
(Mentor 1961).

[23] Id.

cratic values and rigorous application of the rule of law, it cannot be said, uncontroversially, that our polity is committed to these values without exception. Hence, the question is not just a factual one, "Do the canons make the law more predictable (or objective or democratic)?" Instead, the issue is a mix of normative and factual considerations, "What is the optimal level of predictability (or objectivity or democracy), and do the canons contribute toward that level?" Stated that way, the issue is all but impossible to test empirically.

3. Openly Normative Justifications

Most of the substantive canons are hard if not impossible to defend on ordinary-use-of-language or this-is-what-the-legislature-would-want grounds. Legislators inveigh against the rule of lenity when they think about it at all; historically, some states' legislatures have adopted statutes precluding application of the rule of lenity in criminal cases.[24] It is more likely that ordinary speakers would expect their punitive directives to be applied with lenity, but popular enthusiasm for strict punishment of criminal lawbreakers renders that conclusion speculative.

In any event, most of the substantive canons might be defended by reference to normative criteria. For example, the rule of lenity is best justified by libertarian and institutional norms: Legislatures are prone to deploy the moral condemnation of the criminal sanction too readily, and the judiciary is the best actor to curtail this process and at least trim its draconian margins, especially when a person's physical liberty is at stake. These justifications might be debated on normative grounds,[25] but factual studies might contribute to the normative debate. Three key issues are:

-   Do legislatures overutilize the criminal sanction?
-   Does the rule of lenity ameliorate some of this overutilization?
-   ˙Does the rule do so at an acceptable cost?

All three issues have obvious normative components— "overutilize" and "ameliorate some" and "acceptable cost"—and pose enormous methodological problems for any effort at rigorous

---

[24] See Livingston Hall, *Strict or Liberal Construction of Penal Statutes*, 48 Harv L Rev 748, 752-54 (1935) (noting the adoption of several state "liberal construction statutes" designed to prevent courts from construing statutes "strictly," to favor accused criminals).

[25] Compare Eskridge, *Dynamic Statutory Interpretation* at 295-96, 305-06 (cited in note 9), with Dan M. Kahan, *Lenity and Federal Common Law Crimes*, 1994 S Ct Rev 345.

empirical study. On the other hand, all three issues do pose fac-
tual questions that might be illuminated by a more casual em-
piricism, including accounts by legal historians and narrativists
as well as legal sociologists and economists:

- Do legislatures often criminalize conduct later considered
  to have been harmless? Is the process of criminalization
  prone to hasty decisions later regretted? Does the legisla-
  ture tend to overreact with draconian penalties?[26]
- Does the Supreme Court (or a state high court) apply the
  rule of lenity in a consistent pattern? If so, does the pat-
  tern serve the constitutional values of the rule? If not, are
  there features of the Court's application that nonetheless
  discourage overcriminalization? Does the legislature re-
  spond when the Court narrows statutes? Do the responses
  serve the goals of the rule?[27]
- Does the Supreme Court (or a state high court) apply the
  rule of lenity in cases where the rule does not serve its un-
  derlying constitutional values? If so, does the legislature
  override the Court most often in those kinds of cases? And
  does the Court learn something from those overrides?

Admittedly, the inquiries implicated in the foregoing questions
are laden with normative judgments about what is "hasty" and
what is "harmless," for example. What casual empiricism can do,
and perhaps the most it can do, is provide a more informed basis for
judgment about these issues. Casual empiricism also acknowledges
the way law, like other crafts, is experiential and context-bound.

Return to the question Cass Sunstein and I have posed for
formalist theories of statutory interpretation generally and the
new textualism in particular.[28] Existing practice and most theory
emphasize both text and statutory context, including legislative
history scorned by the new textualists. For most of us pragma-
tists, the burden is on the new textualists to justify the signifi-
cant changes in judicial practice their theory demands. The new

---

[26] For a case study of how legislators and prosecutors overreacted to the visibility of
gender-bending and gay subcultures with "excessive" criminal sanctions, see William N.
Eskridge, Jr., *Privacy Jurisprudence and the Apartheid of the Closet, 1946-1961*, 24 Fla St
U L Rev 703, 710-33 (1997). Compare id at 777-81 (describing how the rule of lenity was
deployed, or not, in various jurisdictions).

[27] See Kahan, 1994 S Ct Rev 345 (cited in note 25) (arguing that the rule of lenity is
not consistently enforced by the Supreme Court and that it does not well serve its most
likely constitutional goal of nondelegation); William N. Eskridge, Jr., *Overriding Supreme
Court Statutory Interpretation Decisions*, 101 Yale L J 331, 348, 351, 362 (1991) (empirical
survey finding that Congress overrides Supreme Court decisions invoking the rule of len-
ity with a greatly disproportionate frequency).

[28] See sources cited in note 1.

textualists would respond that the burden is on contextualists to justify departure from what they consider constitutional base-lines. The debate is now at impasse. Would empirical studies break the logjam?

Not in the short or medium term. Just as democracy or rule of law defenses of the *inclusio unius* and the deference to agency canons (new textualist favorites, by the way) would be hard to establish empirically, so would the new textualists be hard pressed to establish empirical defenses for their broader claims. The reason is a general one. Testing theories of interpretation empirically is difficult because what you're testing is hard to quantify, the human and institutional variables are numerous and hard to sort, and unconscious normative assumptions can bias the sampling and evaluation processes. The application of statutes to facts is too spongy to test as rigorously as other things. In the end, the tests may not matter because of people's preexisting normative commitments.

In the longer term, however, empirical testing of textualism and other theories of statutory interpretation ought to have stronger consequences. The strongest consequences would likely be byproducts rather than goals of Chicago-style empirical testing. Because testing would have to be collaborative, new and increased linkages would be formed between law professors and academics in the social and mathematical sciences. These linkages would probably have a variety of consequences, perhaps downscaling the theoretical aspirations of law professors, regularizing tenure or professional standards, and maybe even increasing judicial and lawyerly interest in legal academics' work product. Some would view these consequences as positive cash-values of an empirical testing project. Others would see them as cash-deficits, to the extent they submerged deep normative issues in number-crunching and discussions of craft.