Arbitrariness review of agency rulemakings has long set “political” influences aside as a special case worthy of special scrutiny. This essay argues that the orthodox account of arbitrariness review in this vein makes some untenable assumptions about both reviewing courts and agencies as agents. If we seek more agency responsiveness to reason rightly defined, then reviewing courts must begin devoting more (scarce) cognitive resources to the monitoring of agencies’ behaviors over time. Reviewing courts should encourage agencies to organize themselves in order to learn-by-doing. This will probably entail paying less attention to the separation of law from fact, science from politics, and judgment from justification.

I first argue that agencies are essentially expert systems, composited together from the spectrum of epistemic domains and professional traditions, leaving them uniquely dependent on the very “political” actors who cycle in and out and bring their “political” biases with them as they come and go. Second, I argue that reviewing courts have come to define their own tasks in arbitrariness review in such fantastically ambitious terms that no mortal could possibly discharge those tasks competently. Because of the adversarial processes by which our courts operate, the predicament is made worse as judges must construct law out of the complex statutory materials of today at the same time they must scrutinize the work of ostensibly expert organizations, all while being forbidden from trusting fully the submissions of their litigants. If a rational agency adapts its own behavior to the known biases of reviewing courts, agency learning will simply exacerbate this predicament still further.

If we seek more agency responsiveness to reason rightly defined, then I argue that reviewing courts must begin devoting more (scarce) cognitive
resources to the monitoring of agencies’ behaviors over time. In particular, reviewing courts should encourage agencies to organize themselves in order to learn-by-doing. This will probably entail paying less attention to the separation of law from fact, science from politics, and judgment from justification. Finally, I argue that those separations are at best matters of degree (and probably pure fiction) and have not served us well, all things considered.

We routinely talk of “policy” and “political” influences in the making of administrative regulations. Agency judgments that include such considerations have long been a special case in administrative law. But what role do these influences really play in agency judgment? Administrative rulemakings today are subjected to external forces of many kinds, some laying emphasis on the supposedly “rational” process of summing the relevant reasons and some laying emphasis on the action’s wider public salience. And an agency’s decision-making process, whatever its actual character, is usually held in contrast to the more “political” moments in government where winners reward the loyal for their loyalty, say, or act out of their own venality. In this paper, I shall argue that separating these two ideal types of decision-making for purposes of judicial review of agency rulemakings is much harder than the conventional thinking assumes. Most of the empirical work on these questions is predicated on this basic misjudgment and most courts’ professed attitudes about judicial review rest upon it as well. Courts and their modes of review, I shall argue, cannot achieve a good perspective on agency decisions in rulemakings. Thus, they are poor judges of agency reasonableness, at least if by that notion we mean agencies’ responsiveness to reason rightly defined. Indeed, if what they mean to be doing is identifying and rooting out “political” decisions, court cases of review may themselves be inadequately responsive to reason, all things considered. Although the strongest form of this objection would undermine judicial review wholesale, I shall defend a qualified version motivated by a pragmatic theory of judicial competence.

At their best, agencies take an iterative approach to the public problems they must address, and they organize themselves to learn-by-doing. Contemporary pragmatists have sought to improve and to make more rigorous agencies’ step-wise use of such methods. But if it is more of their “experimentalism” in the judicial review of agency rulemakings that we seek, judicial review should be conducted more mindfully of the court’s own influence and limitations. Doing so, I argue, would lead individual judges to bracket the particulars of a case and shift more of their attention to the ways in which agency behavior and public problem-solving can be integrated and thereby justified over time.

John Dewey observed that democracy is “organized intelligence” and that the single criterion for judging among different models of democracy is “the degree of organization of the public which is attained.” Opinion is still divided over how or what Dewey conceived of as the “public,” certainly, but one thing is obviously not what Dewey thought of as public: the singular, partial, partisan, or personal motivations – undimmed by the needs or claims of others. Dewey’s
own solution to the problematic notion of publicity was to pluralize the public, remove it from its dualistic opposition with the “private,” and to reconstruct publicness as the fuller illumination of all extant interests, motives, and concerns. Thus, Dewey’s public is public by being communicative, deliberate, and developmental. The essay is in four parts. Section 1 introduces judicial review and the function of standards of review therein. Section 2 raises prudential and epistemic worries about the practice of arbitrariness review in court. Section 3 introduces the “experimentalist” therapies proposed by Dorf and Sabel in 1998, and Section 4 suggests a few ways to sharpen and improve the experimentalists’ program.

1.

Section 10(e) of the Administrative Procedure Act (APA) of 1946 set a broadly applicable default that still governs the judicial review of most administrative agency “action.” As codified, it states first that “[t]o the extent necessary to decision and when presented, the reviewing court shall decide all relevant questions of law, interpret constitutional and statutory provisions, and determine the meaning or applicability of the terms of an agency action.” Following that mandate, the APA supplies six separate grounds upon which a reviewing court is empowered to “hold unlawful and set aside agency action, findings, and conclusions.” Finally, following the specified grounds for legal relief, the APA states that “[i]n making the foregoing determinations, the court shall review the whole record or those parts of it cited by a party, and due account shall be taken of the rule of prejudicial error.” The scope, meaning, and inter-play of these different grounds for reversal and other “review” of agency action have dominated contemporary administrative law like no other set of issues. And one form of review has done so more than the others: the reversal of agency action deemed to be “arbitrary” or unreasoned. Shed of its connotations of tyranny, official arbitrariness comes down to some lack of reason or justification. We may even, for the sake of argument, call it the opposite of “reasoned decision-making.” This section explains these standards of review as they co-evolved in our tradition with special attention to arbitrariness review and its demand for reasons and reasoned decision-making.

Long before the APA, administrative agencies like the Interstate Commerce Commission pursued their goals in what we would recognize as an “adjudicative” mode. They summoned and took testimony, compiled evidence, and maintained a formal record, all in preparation for an eventual directive or permission of some kind. It was directed or issued to specific – usually named – parties on grounds particularized to them. And it was proceedings of this kind that generated the original doctrines of administrative law, doctrines that preceded and shaped the APA. Cases that predate the APA are still cited as precedent in federal court today. Indeed, the statutes, precedents and commentary that had coalesced and given form to the APA’s texts were only very gradually replaced
by the APA itself as the locus of controlling legal authority. For much of the APA’s history, indeed, its relationship to the underlying federal “common law” bearing on the very same legal issues it treats has been deeply unclear. Were the APA’s texts intended to be exclusive, comprehensive, and mandatory, or rather permissive, advisory, and interstitial? Inherent judicial authority regarding jurisdiction, choice of remedy, and timing of proceedings long encouraged – and still permits – departure from the APA’s seemingly permanent texts. With administrative agencies engaged in general and prospective rulemaking as much as they are today, however, this ever-present threat of departure from statutory text goes to the very core of the court/agency relationship in contemporary administrative law.

Before they appeared in judicial review of agency action, different “standards” of review had evolved to reflect our popular sovereignty’s peculiar institutional legacy. The vertical and horizontal boundaries of our hierarchical court systems, our separate “branches,” and our Seventh Amendment promoted their development. Within the judicial hierarchy, the principal distinction that emerged was between the review of “findings of fact” and the review of “conclusions of law.” This fact/law dichotomy first grew out of our Constitution’s entrenchment of juries and the jury’s common law authority in the courts of law. Federal appellate courts eventually fashioned two standards of review: the “de novo” standard for questions of law and the “substantial evidence” and “clearly erroneous” standard for questions of fact. Where the former was akin to a complete slate-cleaning (as its Latin origins suggest), the latter was more constrained by and bound to the original decision. Significantly, the standard of review for jury findings has long imagined a “rational” or “reasonable” person who must hear and weigh competing evidentiary claims. Thus, the “clearly erroneous” label has been interpreted to require a presumption of correctness, a refusal to “duplicate” the role of the original decision-maker, or just simply a kind of “deference” to the fact-finder’s better proximity in such matters than that of a reviewing court. Each formulation varies subtly from the others. None results in an affirmance, however, if “no reasonable person” could have drawn the conclusion(s) from the evidence under review. Indeed, the courts regularly expound on the standard in ways that emphasize unmistakably that persons are making the findings of fact. Persons, however, are notoriously motivated – sometimes consciously, sometimes unconsciously – and those motivations are notoriously opaque. Personal motivations are the antithesis of reason, whether it is because they stem from bias or something worse. Not all of them are nefarious, of course, but a “bias” is by definition a departure from or lack of reason. Courts have long focused here, if only tacitly, in reviewing decisions within the judicial system. Thus, for example, while the standard supposedly “does not change,” “the likelihood that the appellate court will rely on the presumption [or correctness] tends to increase when trial judges have lived with the controversy for weeks or months instead of just a few hours.”
motivated thinking in this context – the opposite of reason – is always something lurking just beneath the surface of any given reason(s).

Finally, a third standard arose to review the inevitable exercises of judgment in the trial and hearing of an entire case or controversy. Such “supervision of litigation” decisions admit of more than one correct answer and often involve an indeterminate range of considerations. Thus, the standard became known as the “abuse of discretion” standard. While the standard obviously assumes two very different properties – that of having discretion and that of exceeding one’s discretion – it has defied precise explication. One court recently offered that an abuse of discretion occurs “(1) when a relevant factor that should have been given significant weight is not considered; (2) when an irrelevant or improper factor is considered and given significant weight; and (3) when all proper factors, and no improper ones, are considered, but the court, in weighing those factors, commits a clear error of judgment.” The abuse of discretion standard so described is about identifying the relevant “factors” of choice and assigning the right weight or force to each in the decision(s) made.

How might we organize these different standards? “Deference” to lower courts first arose as to findings of fact and exercises of discretion. Deference like this traditionally meant that reviewing courts adopted a so-called presumption of validity, i.e., that they refused to reverse absent a substantial reason for doing so. But the relative strength of this presumption was never settled, nor was the minimum necessary strength of the reasons that could rebut it. Indeed, well before the APA entrenched it, courts had settled on a consensus that findings of fact, both by lower courts and by administrative agencies, need only be supported by “substantial evidence,” not all of it. By the 1930s, an initial formulation – requiring but a “scintilla” of supporting evidence – had been rejected but not replaced. “Substantial” evidence was more than a scintilla, less than everything in the record, but otherwise equivocal in quantity. By the time of the APA’s enactment, reviewing courts knew that they could not simply presume the facts to support an agency action under review the same way they could in assessing legislative action. It was regarded as a burden to support an agency action with substantial evidence. The more Congress used the phrase ‘substantial evidence’ in its statutes concerning the review of administrative agency action or the judgments of specialized courts, the more routine this supposedly deferential form of review grew. Following its codification in the APA, reviewing courts came to regard the substantial evidence standard as the default for fact review even as its precise meaning continued to evade capture.

Throughout, conduct of a substantial evidence review remained linked to agency proceedings with a defined record. If an agency had elected not to proceed “on the record,” review of that agency’s action in search of substantial evidence remained out of order unless specifically required by statute. The default requirements for proceedings on the record (the APA’s Sections 7 and 8) are triggered only by enabling statutes that require the agency to proceed by “hearing on the record.” In practice, without deliberately taking evidence into a
record sorted as such, an agency is unlikely to meet the expectations of substantial evidence review. The bare notion of evidence, after all, taps into a complex common law tradition wherein proffered information must be prepared and ordered logically, its relevance weighed, and its inclusion or exclusion from deliberations based upon its expected contributory value. Substantial evidence assumes that what is being scrutinized is probative knowledge instead of noise. Challenged propositions must be proven one by one in that tradition. The legal case itself is what decomposes matters to some number of elemental disputes of fact – some number of propositions. Judging in that vein supposedly can single out the motivated or biased.

Administrative law’s combining of traditions has led to some significant doubt about how “substantial evidence” review compares to arbitrariness review of factual findings, though. Is one less deferential than the other? One court that squarely confronted the practical difference(s) between substantial evidence review and Section 10(e)’s arbitrariness standard applicable to issues of fact in “off-the-record” proceedings concluded that there was no substantive difference between the two standards. The only difference, then-judge Scalia held, is that the factual support for agency action under the APA’s “substantial evidence” review must “be found within the record of closed-record proceedings to which it exclusively applies.” Assuming this is correct, it can still make a big practical difference. The APA’s requirement that the court “review the whole record or those parts of it cited by a party” invites the agency’s opponent(s) to submit, and to argue the probative value of, evidence contrary to the agency’s findings. The judicial code affords some flexibility as to what exactly must be filed with the court in response to a petition for review. But review that must rely exclusively upon “record evidence” is more confined than one that permits the use of information from beyond the proceedings to justify or to question the challenged action. It also undoubtedly empowers some within the agency and not others.

Finally, as the Supreme Court has made clear several times over, agencies are not to be held to procedural requirements like keeping a record unless the governing law – fairly interpreted – levies them. Agencies are free to proceed “informally” and without a “closed” record unless a statute or their own rules require it, not least because the closed record proceedings also allocate a burden of proof to the “proponent” of the agency’s action. Thus, a converse holds as well. In instances of “informal,” open-record proceedings, agency action is to be reviewed according to the ‘record’ – such as it may be – actually accumulated and filed by the agency.

A reviewing court convinced that it must serve some kind of checking or other “supervisory” function may expect discrete and identifiable “findings” of fact. And this role-derived expectation attaches regardless of agency action-type. That court will seek to scrutinize the agency’s explanation of its policy-making choices. And it will only defer to the agency’s interpretations of law when the doctrines of deference thereon – doctrines like Chevron and its progeny –
require or permit that deference. This kind of probing examination of agency action has cemented a synthetic tradition in place that we might call the \textit{standard deduction}. It is standard because it holds across contexts in the administrative process, regardless of statute, agency, or action-type. And it is deductive in logical form, at least in part. This standard deduction has two parts:

\begin{enumerate}
\item Because Congress, in delegating the authority it did to the agency, was relying upon that agency’s greater knowledge, expertise, field of view, span of control, \etc, no reviewing court is free simply to substitute its faculties of perception, inference, judgment, \etc, for that of the agency;
\item provided that the agency enables a reviewing court to discharge its checking function, whether by the filing of a “record,” the issuance of a satisfactory explanation, and/or the announcement of discrete legal conclusions, respecting Congress’s choice to afford those affected their right(s) to judicial review.
\end{enumerate}

While this standard deduction has an air of invincibility today, it is open to several doubts. Consider how it has worked in practice. Reviewing courts have obviously interpreted their checking function quite variably. Moreover, the fact/law dichotomy is only as good as the distinction on which it rests. Clearly it does not exhaust what a reviewing court might find fault with in any given agency “action.” The APA commands a reviewing court to “hold unlawful and set aside” three distinct things: “agency action, findings, and conclusions” where any of them is “found to be ... arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law.” Now, as Dewey’s work long emphasized, dualisms like law and fact purport to exclude a middle that almost certainly exists in the world as we experience it. Indeed, factual and legal issues rarely fit into neatly separated compartments in litigated cases, \ie, the cases the opponents are all at least somewhat confident they can win. Most cases that are close enough to sustain themselves in litigation present issues that are mixtures of law and fact. Indeed, the Supreme Court recognized this long ago as a reason to defer as much to agencies’ legal conclusions as to their findings of fact.\footnote{Mixed questions of law and fact arise often enough, but even where law and fact are readily distinguished there is usually some residue left over within the agency’s “action” for a reviewing court to scrutinize. \textit{Action} has always implied a totality that is broader in scope than any particular intentional state. Philosophers of action must explain intention, consequence, knowledge, motive, and the like.\footnote{Action in this context expands what may be scrutinized. To take an all too familiar example, where the governing law posits a range of goals or priorities for the agency to sort, law, fact, and judgment can intertwine with one another so as to necessitate review of an overall explanation that falls under no particularly defined standard of review.\footnote{This is why the “hard look” standard}}
of review has always aimed for something beyond the traditional law/fact
dichotomy. But its exact aim has shifted over time. As is often said, the concern
is the agency’s reasons and, more specifically, some sort of assurance that the
agency action under review is the result of “reasoned” decision-making.  
Reasons for action have turned out to be a troubling focus in the court-agency
relationship, though. The trouble is how little we can say about reasons for
agency action which does not require, for its compatibility with our Constitu-
tion, that we empty from the notion of reasons for action all that makes the
notion demonstrative, explanatory, or communicative.

Reasons for agency action have a surprisingly complicated structure
which judges engaged in judicial review of agency action probably cannot
effectively scrutinize or improve. The first cases from the D.C. Circuit to
mention a “hard look” mentioned it as what the reviewing court expected the
agency to take to the facts and issues before it.  
Judge Leventhal, its original
proponent, said that the standard “calls for insistence that the agency articulate
with reasonable clarity its reasons for decision, and identify the significance of
the crucial facts, a course that tends to assure that the agency’s policies
effectuate general standards applied without unreasonable discrimination.”
Shorn of its embroidery, Leventhal’s hard look doctrine demanded perspicuity
of the agency and, above all, a willingness to hear the arguments and to weigh
the options presented by the agency’s opponents.

What we are entitled to at all events is a careful identification by the
Secretary, when his proposed standards are challenged, of the reasons
why he chooses to follow one course rather than another. Where that
choice purports to be based on the existence of certain determinable fact,
the Secretary must, in form as well as substance, find those facts from
evidence in the record. By the same token, when the Secretary is obliged
to make policy judgments where no factual certainties exist or where
facts alone do not provide the answer, he should so state and go on to
identify the considerations he found persuasive.

For a time, the accountability Leventhal envisioned here between agency
and reviewing court seemed productive and healthy. But it was short-lived.
The more insistent the courts grew that agencies offer perspicuous explanations
for their actions, the more the (motivated) litigants bringing their challenges
learned to find gaps or other weaknesses in agencies’ given reasons. Motivations
and biases are, after all, equally a property of litigants who go to court as they
are of agency decision-makers. Moreover, the formative precedents on this hard
look review examined law, fact, and judgment as they intertwined in open- and
closed-record proceedings alike. Consequently, agency actions eventually
became so indirectly “proceduralized” by the D.C. Circuit’s hard look that the
Supreme Court intervened and not-so-politely admonished the lower federal
courts that “administrative agencies should be free to fashion their own rules of
procedure and to pursue methods of inquiry capable of permitting them to discharge their multitudinous duties. The more complex and technical the issues grew, the less confident the judges themselves felt in their own competence to assess an agency’s reasons from a bare record of that agency’s proceedings. The more frequently complex factual records came to court as the basis for some agency’s rule, the more uneven the review became. Section 2 uses one recent example of arbitrariness review in this tradition to illustrate its paradoxical character.

2.

In theory, agencies can organize and operate to make much fuller use of domain experts than can the courts. This is the comparative advantage of bureaucracy over less rational institutional forms. In practice, however, they can do so – consistent with the standard deduction – only if they can also explain their methods and choices to the reviewing court’s satisfaction. When the Supreme Court finally defined its version of arbitrariness review, it began with a restatement of the standard deduction. And then it opined that

an agency rule would be arbitrary and capricious if the agency relied on factors which Congress had not intended it to consider, entirely failed to consider an important aspect of the problem, offered an explanation for its decision that runs counter to the evidence before the agency or is so implausible that it could not be ascribed to a difference in view or the product of agency expertise.

This formula in State Farm has since taken over the mantle of the “hard look.” It presupposes at least the following: (1) that agency judgment consists in identifying and (properly) weighting the discrete “factors” of choice; and (2) that the agency is under a duty to substantiate a factual basis for any judgment(s) that can accord with any express justification(s) sufficiently well for a reviewing court to endorse. Each of these commitments is substantive and problematic. Let us take the second one first.

At issue in State Farm was a National Highway Traffic Safety Administration’s (NHTSA) standard on automobile passenger restraints under the Automobile Safety Act – an agency action Congress supposedly required be tested against a “substantial evidence” standard. In a proviso that has inexplicably gone missing ever since, the State Farm Court also held that “Congress required a record of the rulemaking proceedings to be compiled and submitted to a reviewing court ... and intended that agency findings under the Act would be supported by ‘substantial evidence on the record considered as a whole.’” And such a requirement followed (if at all) from NHTSA’s unique enabling statute – not the APA. Much of what animated the Court’s actual scrutiny in State Farm was the extensive factual record NHTSA had amassed to
support its challenged action. Thus, the scope of State Farm’s arbitrariness review stemmed – ironically enough – from a “substantial evidence” agency proceeding. Clearly, NHTSA’s “burden of explaining” its course of action in the convoluted history of that particular standard-setting was heightened by the presence of so much record evidence and documented agency indecision. And this nexus in State Farm was probably uniquely demonstrative to agency lawyers: the more the agency documented the more it had to explain, especially because so much of what was being explained was the range of expert opinion. The economists’ counsel, when combined with the crash test experts’ counsel, may have been easily recordable. But they were seamless neither in the direction nor in the degree of action for which they served as counsel.

The prevailing process norms in rulemakings governed by APA Section 4 demand that agencies put out for public scrutiny whatever information, data or other evidence they believe supports their proposed rule. This, in turn, typically elicits responsive submissions and commentary. But what is responsive is heavily influenced by the initial framing of the rulemaking proposal. Without any bars on admissibility (as the rules of evidence empower trial judges to erect, especially as to expert opinion), the records from this process can be rather hard to conceptualize. What factual predicate could possibly be amassed and transmitted to justify a rule that embodies a compromise between two mutually competitive projects like risk aversion and wealth maximization? Normally, different objectives like this will involve (at least) two different kinds of domain experts, demanding that the agency’s precise mediation between the two be substantiated factually. Without knowing what the standard(s) of proof should be for such intermediation, the “records” being compiled will be ad hoc. Furthermore, the non-expert judges reviewing such records, any constituent findings, and/or any attendant explanations, are in a precarious position. Whatever biases may or may not affect agency personnel, without a formal record, reviewing courts are hardly in position to establish it after the fact. And though they cannot impose the judicial system’s traditions of evidentiary proof, using any other identifiable standard by which to assess an agency’s factual work would arguably encroach upon the constitutional role of the legislative and/or executive branches. Deliberately creating a standard could seem either like making law or assuming the President’s duty to “take care that the laws be faithfully executed.”

Consider a recent example and how it problematizes both of State Farm’s presuppositions in the review of agency reasoning. In Massachusetts v. EPA, the Supreme Court reviewed an agency action having both tremendous political stakes and notoriety and unsurpassed factual complexities. EPA was petitioned to reach a finding under Clean Air Act § 202(a)(1) that “greenhouse gas emissions from new motor vehicles” are “air pollutants” the emission of which “cause, or contribute to, air pollution which may reasonably be anticipated to endanger public health or welfare.” The agency rejected the petition and review was sought. In a 5–4 decision most regarded as a pointed rejection of the
Administration’s political involvement in EPA’s decision, the Court held that the greenhouse gas emissions complained of could not but be “air pollutants” within the meaning of the Act and that the only factual determination left for the agency was whether the emissions “cause or contribute to” endangerment of the American public. The Act’s definition of “air pollutant” is designedly broad. And, as the majority implied with its review of various institutions’ responses to the mounting (scientific) evidence proving global climate disruption, assuming greenhouse gases are “pollution,” they are probably the worst pollution problem we will ever face.

Nowhere in the informal “record” compiled, though, did EPA ever assert or even imply anything to the contrary. The agency never said that greenhouse gases do not or cannot endanger public health or welfare. The agency’s conclusion was that Congress had not adequately “authorize[d] regulation under the [Clean Air Act] to address global climate change.” All of EPA’s reasons for denying the petition were to that effect, chief among them the agency’s interpretation of various other statutes bearing on global climate disruption (none of which empowered EPA to regulate emissions) and the agency’s familiarity with the Clean Air Act’s legislative history. Taking its cue from the Supreme Court itself, EPA compared the scale and scope of the problem of global climate disruption against the agency’s authority and decided that the Act’s general grant of authority on this problem of unprecedented scope—which Congress had several times considered but for which it had never delegated regulatory authority expressly—was insufficient. EPA rejected a rulemaking petition, that is, it did not “find” that “greenhouse gases” are not “air pollutants.” In its own words, EPA’s “action” was not a finding regarding the polluting character of greenhouse gas emissions. The agency explicitly avoided making such a finding, noting repeatedly that although the statute mandated an up-or-down decision on the petitioned compounds as “air pollutants” under the Act, it contained nothing like a deadline for that particular decision/action. The Act did not bind EPA to move within any certain interval of time. Thus, in EPA’s view, while it continued to consider the problem (as would everyone else, including Congress), it could hold off from regulating what amounted to a minute fraction of that problem.

Confronted with the agency’s statement of reasons and an apparent sense that the President’s politics had motivated EPA, a majority of the Court held that EPA could not rightly consider the causal redundancy of greenhouse gas-forced global warming in this way. EPA’s assertion that emissions of these gases around the planet were going to endanger public health and welfare regardless of future U.S. auto emissions was highlighted by the majority as an unauthorized reason for denying the petition. It was, in State Farm’s formula, a “factor[] which Congress had not intended it to consider.” But how could this be so? How can the fact—or at least EPA’s supposed finding—that global warming is causally over-determined be completely excluded from EPA’s decision whether to regulate emissions from (certain) U.S. sources as an “endangerment”? In our
legal traditions, additional causes are always relevant and never more so than when the additional causes would themselves be sufficient to bring about the harms in question.\textsuperscript{92} We care about causes because we care about what links behavior to its consequences.\textsuperscript{93} Things would have to turn upside down for an agency of ours to ignore completely a causal redundancy it noticed in greenhouse gas emissions the planet over. At the very least, the statute should have been explicit to that effect, and it was not.

\textit{Massachusetts} arguably highlights a more general failing of the Court’s in applying \textit{State Farm}. How could the supposed exclusion of EPA’s reasons found by the \textit{Massachusetts} majority have been viewed as some kind of “clear statutory command?”\textsuperscript{94} Could the statute really have meant EPA to ignore the quantities of the pollution as causes of the pollution problem?\textsuperscript{95} From all indications in prior Clean Air Act endangerment precedents, EPA’s disregard of a causation failure like the one it identified would itself have been regarded as “arbitrary.”\textsuperscript{96} The paradox is obvious and familiar to lawyers: the statutory constraints supposedly structuring agency judgments of this kind are so plastic that they are manipulable. Modern statutes are hardly constraints at all in the kinds of decisional actions so often challenged in arbitrariness review. Their strictures amount to a check on agency carelessness more than they do a check on lawlessness. The Supreme Court has repeatedly said as much about the statutes that create these contexts – refusing to scrutinize them for anything like a single or uniquely correct meaning when the issues raised are both specific and programmatic.\textsuperscript{97} \textit{State Farm}’s very notion of agency choice consisting in some isolation and proper weighting of choice “factors” depends on the opposite presumption, though. It naïvely presumes a discernible set and, by implication, the exclusion of non-members. But sets that are both tractable and exact are rare unless they are stipulated in explicit language. Indeed, sets are more often paradoxical.\textsuperscript{98} Without some ground on which to stand to find and weigh choice factors as the members of a \textit{State Farm} set, the actors are left with nothing but space.

Take the Clean Air Act’s proviso that pollutants need only “contribute to”\textsuperscript{99} the pollution problems at issue for EPA to make an endangerment finding. Put aside the dissent’s argument that the phrase is “ambiguous” in the sense that the Court had previously held entitled EPA’s interpretation to judicial deference.\textsuperscript{100} What ought to be a relevant contribution where air pollution is concerned? Contributions like this are themselves paradoxical: what is enough to be relevant? It cannot be enough to say anything no matter how negligible, for that just negates the significance of EPA’s decision and, along with it, whatever sort of functionality or accountability we seek of agencies like EPA. So where is the threshold? The troubles compound when we remember that the greenhouse effect threatens us environmentally in a way unlike any other “pollution” hazard to which the Clean Air Act has been applied. Assuming no particular quantitative threshold is more easily inferred from the Act than any other, how should it be decided? New motor vehicle emissions in the United States over the coming
decades will be, under any reasonable estimate, no more than a tiny fraction of
the causal forcing behind the greenhouse effect. We could call this our
delegation problem: if statutes are going to be so general as to accommodate the
unforeseen, when the unforeseen unfolds, it may well challenge all of our
original suppositions – leaving the administrators of these statutes in the
unenviable position of having to contort and/or uncontort them to find a
workable fit.

Contemporary statutes and their interpretation add more than just a layer
of uncertainty to such problems in practical reason. The delegation problem is
pervasive. Such statutes make State Farm’s notion of a choice “factor” into a
drag on practical reason when agencies confront the unforeseen by com-
ounding questions of constitutional role and linguistic convention/logic with
questions of causality, consequence, and epistemic competence. Identifying
choice factors from statutes at decisional junctures like the one at issue in
Massachusetts – which is to say, highly specified junctures where subtle
variations can be decisive – normally presents the very kinds of hard questions
that judges work hardest to avoid. Their ready alternative is the mere suggestion
that the agency has considered something it should not have or failed adequately
to consider that which it must. The ready alternative is the implication that an
expert has fallen prey to motivated or biased reasoning – like some jury or errant
judge.

When courts wish to forbid, permit, or require an agency’s attention to
some “consideration,” they label it a “factor” within the State Farm formula for
reviewing agency judgment. These “factors” can be derived from statutory
text, legislative history, precedent, or agency practice. Just as the hard
look doctrine was emerging, several cases involving the setting of environ-
mental quality standards by agencies like OSHA and EPA were litigated where
statutory factors became the principal question. Whether the agency was
permitted, forbidden, or required to consider the supposed “costs” of environ-
mental quality was a common dispute. Because statutes so rarely address that
question clearly, though, reviewing courts have been left to interpret statutory
“silences,” legislative history and other similarly under-determinative sources.

With so little expressed about what may, must, or must not be a relevant
consideration in an agency’s reason balancing, State Farm reviews that imply
political biasing need never differentiate effective coordination of an administra-
tion’s strategy or priorities from, for example, meddling on behalf of a political
client. Like the scholars who have opined about proper versus improper
“political” influences in this context, reviewing courts need never get any more
specific about the distinction, if there is one.

There may be no Archimedean search for State Farm’s factors, but
combining the trappings of litigation with the background levels of uncertainty
inherent in statutory analysis – to say nothing of the amateur evaluation of actual
science – raises serious epistemic worries. Agencies are supposed to
aggregate different human capacities to achieve that which individual generalists
They are supposed to be “expert systems,” so to speak. This is hard to reconcile with the idea of a generalist judge’s reviewing such an organization’s work. Whether an “explanation runs counter to the evidence” turns on how one can square the evidence with the explanation(s) offered. Attitudinal influences surely play a part. More perversely, though, agencies, their explanations, and the evidence they gather to justify their judgments are all viewed in the adversarial process by which our courts operate as partisan. They are not to be trusted fully when they are in court. It is the judge’s constitutional duty to take this attitude, the judge’s duty to be “fair-minded” about the facts until the time for his or her conclusions. Yet these decisions are reached by the same persons who must at the same time make sense of a (likely complicated) statute. Combining the two tasks is compounding the challenge of the office. Forcing one decision-maker to construe the law, to review the findings and judgments of a more or less expert organization, and to interpret partisan arguments in order to do it all is expecting a lot. The well-known biases of all reasoners would be daunting to any but the most confident of judges.

These failures of reason in the courts may be of broader significance in themselves for the simple reason that agency lawyers are aware of them. For example, is the judge inclined to view the statute as picking out genuine kinds through some referential history or is s/he inclined to view the statute in a more cognitivist vein whereby the legislators’ collective intentions matter most? Are these inclinations strong enough to persist from case to case? If not, that all by itself may undermine the normativity of whatever “factors” for agency judgment hard look review can select. After all, if judges have no robust inclinations on this point, they may be moved by reasons that end up being mutually inconsistent in short order. Agencies making rules must prepare to deal with “the judiciary,” a fictional whole standing in for what is, in truth, a (somewhat) diverse population of individuals. But if agency lawyers know that the “arguments” can be changed-up whatever the selected judges’ past judgments, it is hard to find anything principled in the statutory interpretations of such a system. Given how formative legacy precedents can be, it is arguably the agency lawyer’s duty to craft winning arguments – not to communicate forthrightly. Whatever plausible “factors” emerge over time in that nexus, in short, will obviously skew away from what a randomized, non-sequential case ordering would create. They will be, in essence, accidental.

Finally, it seems wrong to assess real expertise as do reviewing courts in this doctrine. It makes the double mistake of viewing agency-gathered evidence and expert advice as too weak and too strong at the same time. In court, every factual dispute has at least “two sides” that must be allowed their hearing. But this oblique approach has less to do with truth or reason than with orthodox notions of fairness and procedural justice. Fact finding in the legal process conceptualizes information in this way. In the world-at-large, we know to be wary of the distorting effects of particulars, though. Facts are rarely so authorita-
tive as the judicial process makes them, of course. So while testing ostensible facts may be the essence of science and modern probability, doing so in the ways that litigation venues allow is deeply problematic. Besides the “anchoring” and “availability” problems inherent in any case-by-case analysis, a professionalized attitude of fairness can find its way into virtually any cognitive space. Yet fairness, as David Estlund has argued, is an occasional value. “It can sometimes seem as if everything should be fair.” Fairness’s hegemony probably has less to do with our actual values, though, than with the accidents of our language. Many things that are not fair are not unfair either. And holding a proposition of fact in abeyance in order to entertain some proposed alternative can only be virtuous, truth-tracking, or reasonable if the judge is qualified to do so. When scientists do so, it is within the confines of their own epistemic domain and at the behest of someone of real standing who has mounted a defined and plausible challenge.

The “implausible” of which State Farm spoke is, we may reasonably worry, precisely the manner of attack that cranks have taken since the emergence of modern science. And the implausible has too often become the truth to take this sitting down. “If we believed that when we consulted our neurologists or our aircraft engineers in the real world there was only a 50/50 chance that they would provide accurate advice, we would not call them experts at all, nor would we pay them handsomely to cure our diseases and build our planes.” Of course, litigation delivers the “expertise” courts must weigh when reviewing agency action in precisely this form. If it is an expert’s words that undergird an agency’s judgment, what is the likelihood that a reviewing court will simply misunderstand those words? After all, “what is plain to those within a specialized linguistic community may be obscure to the man or woman on the street, and the same word or phrase may mean different things to different linguistic communities.” There is at least some propensity of generalist judges simply to misunderstand the expert’s counsel to an agency, probably exacerbated by the ad hoc nature of the “records” conveying that counsel and the agency’s reactions to the particular litigation. Finally, a court that uses State Farm as it was used in Massachusetts – surfacing and rejecting what it thought were illicit “political” influences – creates precedent to do so in the future, whatever the direction or degree of inclinations on future courts.

Try as they have to find an approach that is both content-independent and appropriate to Article III judging, reviewing courts mostly know that their charge in arbitrariness review of agency rulemaking is fundamentally askew. Judges are charged with doing that for which they have no comparative institutional advantage and too little absolute competence.

3.

Section 3 argued that State Farm must assume agency choice takes a form it does not in fact take and that the hard look as State Farm formulated it demands
too much of reviewing courts. The law can forbid, permit, or require conduct but
statutes today are characteristically vague as to what, if anything, they actually
accomplish to that end. Statutes like those that gave us the “standard deduction”
set out in Section 1 normally direct an agency to establish the norms that forbid,
permit, or require conduct and to do so in pursuit of some list of goals or
priorities. Agencies are institutionally complex, though, and they use a variety of
tools to coordinate the inter-operation of their many parts. A common tool is the
rule issued by agency superiors which directs subordinates how to specify the
rights and duties of a regulated party. The question arises how such rules figure
into the reasoning of “the agency” – how the rules interact with otherwise good
and sufficient reason(s) for agency action. This section argues that judges are
in no position to answer this question with much precision and that the hard look
therefore presumes reviewing courts are better off epistemically than they really
are. For, if courts cannot recognize and differentiate the uses of authority by and
within administrative agencies, how agencies actually respond to reason as
opposed to something else will remain obscured no matter how hard the look.

The conventional wisdom is that an agency’s norms are binding on the
agency itself until repealed. Agency norms that bind in this way exclude or
defeat reasons, so to speak, and supply the basis for some agency judgment that
would otherwise be inadequately justified. To say the agency decided as it did in
compliance with its norms is to say that the applicable norm was the agency’s
justification, its reason for action. This is as conventional in contemporary legal
philosophy, e.g., the work of Joseph Raz and other positivists, as it is in
established doctrines of administrative law. To follow a norm is to ignore any
countervailing reasons for action and that is precisely why affected parties may
seek review of the agency norm itself in advance of its application to them. One complication in contemporary administrative law is the APA’s definition of
“rule,” which includes every agency “statement” that is “of future effect” and
which is “designed to implement, interpret, or prescribe law or policy.” This
definition is, in the context of mostly bureaucratic organizations, extraordinarily
broad. Agencies issue “rules” in this sense constantly; it is how they function as
coherent organizations. Thus, the binding force of agency rules in all their
variability and diversity is a matter of some controversy. While all agree that
agency rules can have binding force only if Congress authorizes the promul-
gation of such rules, the coercive force of agency authority can vest ostensibly
nonbinding “internal” rules with this property indirectly. Many agency
“policies” turn out, because of the way the agency is organized, to exclude or
defeat reasons just as a binding norm would. And the necessary and sufficient
aspects of statutes that will authorize agencies to make practically binding
norms like this remain deeply unclear more than a century after the problem
was recognized.

In their 1998 account of “democratic experimentalism,” Dorf and Sabel
argued that judicial review of means/ends rationality had become over-
elaborated in doctrine and verbiage and that it has always been about basic
means-ends analysis. “Stripped of the confounding complexity of means-ends scrutiny in its familiar forms, an experimentalist review moves in the direction of an express jurisprudence of excluded or impermissible reasons.”136 Dorf and Sabel argued that this jurisprudence could replace the dichotomy of today wherein courts oscillate between “extremes of deference and intrusion.”137 But their notion of reasons – much like the Supreme Court’s notion of factors for decision – was vague and perhaps not strong enough to improve agency rationality through judicial review. Dorf and Sabel imagine reviewing courts as trying in earnest to “smoke out” the impermissible reasons – the biases and political motivations, etc. – lying within agency actions. The notion is attractive as many have adopted this view of the hard look and its power to unearth prohibited or otherwise defective reasons. But the nature of agency norms and organization render this more a control on agency carelessness than on the formation of illicit motives by/within administrative agencies. Let us take as an example the Federal Aviation Administration’s 47-year prohibition on commercial pilots flying past their 60th birthday.138

In 1959, the FAA adopted by notice and comment what it termed a safety rule prohibiting regulated air carriers from employing pilots older than 59.139 The ensuing challenges engulfed FAA for virtually the entire half-century of that rule’s validity. The first challenges were that the rule was an arbitrary cut-off and that its application was the denial of an individuated hearing to which every licensed pilot should be entitled.140 Soon thereafter, pilots combined challenges to the rule itself with challenges to their own exemption denials.141 Eventually, pilots were arguing not only that the rule lacked the required means/ends justification but that it had been adopted as a favor to the airline industry as an indirect control on labor costs.142 Each of the successive challenges was rejected by FAA despite evidence that the rule was unjustified. As nine courts of appeal heard and rejected arbitrariness claims, FAA’s reasoning slowly came into focus. FAA had adopted the rule, it said, because the progressive degenerative effects of aging make it increasingly likely that aging pilots will experience some kind of sudden incapacity that is unpredictable in individuals but entirely too predictable across the population.143 Barring some kind of breakthrough that would improve medicine’s ability to predict these events in individuals, thus, FAA held that the rule and the risk put the burden of proof on the individual pilot.144 And that burden was practically insurmountable, not least because no U.S. pilots were flying large planes commercially past 59.145 The agency even adopted a policy of rejecting all exemptions unless and until an applicant presented the medical breakthrough FAA said was needed.146

Notice that the practical reasoning behind the Age 60 rule in 1959 differed markedly from that which carried it through successive arbitrariness reviews. While there never was any medical significance to age 60 as such, the essential issue was that no medically significant threshold could (or can) be set to correspond to the risks FAA was regulating. There was, in some sense, a deficient justification for the agency’s actions. But notice, too, that in challenge
after challenge, the courts never did “smoke out” any impermissible reason for maintaining the rule – and neither did they dispel the doubt that FAA was acting for bad reasons.147 Clearly if the rule were a favor to the airlines, that would be a bad reason.148 But what if the agency was simply responding to the biased or motivated reasoning of those around it? This is where arbitrariness review as traditionally conceived fails us. It offers no conception of the agency as an actor. “Agent specific” reasons are not necessarily bad reasons, but who are the relevant agents in an agency? Standards of judicial review have inevitably boiled down to whether a “reasonable person” could think what the subject thinks. Agencies are a composite of different kinds of people, though, even while the agency as a whole can have interests and concerns unto itself. So what if FAA retained its rule because hard look scrutiny and State Farm would have made any action changing the rule into a struggle the agency preferred to avoid? Justifying any alternative threshold to a reviewing court’s satisfaction in light of the text of the Federal Aviation Act (requiring FAA to regulate “in the interest of safety”), State Farm, and the available information on the nature of aging would have been difficult at least.149 Confronting these “transitional” risks, perhaps FAA decided that exempting experienced/elderly pilots one-by-one would also be a mistake. (Imagine if an FAA-exempted pilot caused an aviation disaster.)

A role-guided hard look of the sort grounded in our standard deduction perhaps represented motivated, biased reasoning to FAA. To FAA, safety in the aviation system is more important than being fair to individual pilots or even the appearance of fairness to all pilots. The role-defining standard deduction introduced in Section 1 even pushed one court to preach in dicta that “it is obvious that the FAA must continue and must enhance its efforts to accommodate [the pilots’] points of view” and that “[o]bviously, there is a great body of opinion that the time has come to move on.”150 Points of view? To an agency, all this communicates is the court’s sympathy for the pilots that almost – not quite, but almost – motivated the court to hold for them. So if FAA thought that trying to improve the Age 60 rule was too risky, could its inaction have been for good and sufficient reasons? How exact would FAA’s tradeoffs have to be to satisfy a State Farm review? Would it have to quantify what it was balancing or would some intuitive weighting suffice? What in the law could authorize FAA to act so strategically? If there is no law authorizing such reasoning, does it represent bias or some other failure of reason? How would the individual pilots’ fights to fly as long as possible151 be factored into a hard look review of FAA’s choice either to keep the rule or to reject all exemption petitions? FAA could reasonably presume that any choice limiting a piloting career (if 65, why not 67?) would be litigated repeatedly and zealously. But could a reviewing court count that as a “factor” for the agency’s decision? And how little consideration of the individual petitions would have been too little?

Finding defective reasoning in the many opportunities presented – or even proving the acceptability of FAA’s reasoning – could only have come from
Reasons as Experiments: Judgment and Justification

223

a better conception of reasons for agency action. Without an adequate account of
the roles that different people should play in an administrative agency,
reviewing courts have to reinvent rational or reasonable agency action over and
over.152 Yet they have no way of knowing who should contribute what to agency
reasoning, when, or how – while at the same time they must ignore their own
biasing effects on agencies as agents. This is our problem. The concept of a
reason is itself relational. “Basic reason statements of the form ‘A has a reason
to Φ’ or ‘here is a reason for A to Φ’ indicate a relation holding between some
agent A and some act Φ (e.g., an action, belief, feeling) – a reason is a reason for
someone to or for something.”153 Our problem is grounded in how little we have
actually decided about agency autonomy – how much of it we want, when,
where, to what effect, etc. This is at least in part because agencies are so diverse
(or at least non-standard), dynamic, and active. But it is also because of our
conflicted attitudes about their roles in society and the legitimacy of reviewing
courts in a democratic order. Section 4 finishes by suggesting a more productive
form of arbitrariness review given our overall predicament.

4.

The standard deduction and the judicial duty to be fair to litigants mostly wipe
the arbitrariness slate clean, iteration to iteration.154 In the two examples
examined, these two have exerted subtle systemic influences, though. How did
the Age 60 rule’s paradoxical justification – hinging on the non-detectability of
age-related deterioration in particular pilots – survive repeatedly and without
experiment for so long? State Farm hard look review allowed the courts to
ignore FAA’s longer-term failings and, therefore, not to prompt any real agency
learning. Without scrutinizing the choices keeping FAA ignorant, each judicial
case was left to its own accidents, so to speak. Powerless to change how the
agency organized itself or who played what role therein, the courts could not but
focus on the “records” brought to them – a matter of some accident across the
range of cases.155 Now Congress eventually increased the mandatory pilot retire-
ment age to 65.156 But was that truly corrective action? Presumably a set of
pilots could fly longer (although that set ended up being slightly deformed by
the process157). But society as a whole simply kicked the problem down the road
a little and learned virtually nothing for its troubles. The agency’s reasons were
never public in Dewey’s sense because they were not communicative,
deliberate, or developmental.

An agency confronting the risk of aging in pilots and the uncertainty of
pilot failure should want to reduce both. It should want to learn whatever can be
learned about both the risks of age-related sudden incapacitation and about the
uncertainty of individual pilot failures. According to an orthodox distinction,
risk is variability that can be expressed in a probability distribution whereas
uncertainty is not so quantifiable.158 On this view, to change your uncertainty
into risk you must learn enough to gather the actuarial knowledge that lay
behind risk expressions as opposed to the paralyzing uncertainty of first encounters. As older pilots argued to FAA time and again, the population of pilots who reached age 65 was actuarially different from the population of all pilots. These were the most experienced people who had never been disqualified, whether by routine medical testing or the hazards of flight. Forcing their retirements may not have been rational at all. Airlines would hire replacements, presumably, and they might actually be more risky than the pilots being forced out. What had FAA done to learn about either the uncertainties of aging or this possible amplification of pilot failure risk? At the very least, improving American life expectancies from 1959 to 2007 (from roughly 69 to roughly 77) should have prompted more FAA experimentation with the relevant norms. An arbitrary age cut-off is a reflection of uncertainty. Doing nothing to further specify the norms governing senior pilot eligibility in almost five decades of practice is a failure to remain communicative, deliberate and developmental. What should a reviewing court have done? If it was evident that arbitrariness review was making the safety of a valid rule into the agency’s crutch, an experimentalist court would have engaged the overall context. A shrewd judge would surely have found some way to fault the agency’s own norms for blocking the gathering of better actuarial data over time. More than some errant dicta to that effect, though, any court could have drawn a connection between its own primitive tools in arbitrariness review as we have institutionalized them and what little FAA was doing to decrease its own arbitrariness. That kind of self-audit may be rare in judicial review today but it would go some distance toward Dewey’s utopia.

In the Clean Air Act greenhouse endangerment case, the Supreme Court’s finding that greenhouse gases from new U.S. autos will be “air pollutants” and its framing of EPA’s 2003 deliberations as demonstrably contrary to the Act was either unreasoned or inadequately communicated. Apparently convinced that EPA’s reasoning was somehow motivated by politics, the Court rejected the agency’s given explanation. But the Court is at least guilty of the same failings it found in EPA: without a good argument that the Clean Air Act cares nothing about causal efficacy in the regulation of “pollution,” the majority’s holding seems motivated or biased. The ensuing (bizarre) EPA rulemakings have all foundered in the very dilemmas EPA first articulated and saw belittled in Massachusetts. Globalized markets limit the practical value of controlling fossil fuels in only some locations, especially because fossil fuel combustion has for so long been so widely identified with freedom and economic growth. Regulatory fragmentation deprives EPA of the jurisdiction needed to control emissions of the six principal (“well-mixing”) greenhouse gases in any economy-wide fashion. The (potentially strategic) reactions of others will be the difference between a pointless drag on one’s own economy and incremental progress toward a carbon-constrained future. The key is in the contributions, and perhaps the contributory ethics, of all relevant others – something no single agency (and surely not one organized like EPA) can dictate. After Massachusetts, EPA now
Reasons as Experiments: Judgment and Justification

seems to have to regulate everything and everyone almost at once even as the United States’ contribution to global climate disruption continues to shrink in relative terms.

Given that the tools in the Act were never designed to solve so en-grossingly global a problem, EPA probably will not find its own best measure of agency autonomy. EPA cannot possibly succeed without the abiding support of the President and a variety of Cabinet-level departments. Indeed, if EPA is mostly a “taker” of the strategies adopted to cope with this enormous collective action problem, the choices of others – merely by changing the relative probabilities of some outcomes – could reorder EPA’s priorities and thus the outward appearances of EPA’s reasoning. Yet, if history is any guide, some of those actors, or those with whom they partner, will also be EPA’s toughest obstacles. These are, in a sense, the wages of the Supreme Court’s naïve separation of science from politics in Massachusetts – and perhaps of reviewing courts that ignore the biasing effects of their own power. There can be nothing illicit in an agency openly confronting the limits of its own agency, its own agential constraints. And with a total of 35 different petitioners and some 80 distinct legal claims challenging the four separate rulemakings that followed from Massachusetts EPA today faces a reason-giving gauntlet of unprecedented proportions. Perhaps now that it is no longer a Bush Administration lamenting the scale or scope of greenhouse gases as “pollution” to be prohibited, but rather an Obama presidency committed to avoiding showdowns that pit the economy versus the environment, the Supreme Court will “defer” to EPA’s reluctance to makes rules without the requisite information. If so, the Court may be able to contain the damage done to practical public reason in our climate politics, at least for the time being. But it will have contributed little to the project.

A pragmatic model of agency arbitrariness starts with agencies as constructed, internally plural, dynamic, and complex selves that evolve over time and depend on their “political” leaders (who bring their own biases) to do so. It avoids dualisms like science versus politics in contexts where they are unavoidably intertwined. And it would probably look beyond the accidents of particular judicial cases to invest (scarce) cognitive resources in monitoring whether an agency is adapting itself in order to learn-by-doing. We need not agree on a complete theory of democracy or on just how agencies can be better agents of the public to agree that they should be organized to learn from experience in context. That low-level principle would have gone a distance toward improving our thinking about climate change had it animated the courts (or EPA) before now. Recording, sorting and publicizing economy-wide emissions data at manageable expense without double counting and over a long enough interval to scrutinize ostensible causes would have been a smart move in 1990. By 2009, it was simply long overdue.

This returns us to the standards of review and their practical differences. If arbitrariness review differs from substantial evidence review because only the latter must proceed from a closed record, that should feedback on and inform a
court’s institutional role in each. If arbitrariness review need not proceed from a closed record, then its signature virtue may be that the predictive and projective techniques so endemic to general rulemaking can be assessed diachronically. That is, with administrative agencies so likely to lose intensity and purpose over time, courts are uniquely positioned to hold them to their promises. If rulemakings must be finalized before all the facts are in, ideally this should result in some kind of “justification I.O.U.,” redeemable in court with the passage of time. For example, if EPA promises that it is not just ignoring most sources by first regulating the enormous ones, its arbitrariness ought to be judged, at least in part, by whether it makes good on that pledge. In the orthodox account, this perhaps boils down to seeking constantly the exchange of uncertainty for risk that is quantifiable, manageable, and distributable. But in a more pragmatic light, reviewing courts view agency reasons as experiments. Experimentation is nothing more than methodical testing with the goal of falsifying, verifying, or validating a hypothesis. They vary in scale and scope, but all experiments rely on transparent, repeatable procedure and analysis after the fact. Reasons for agency action cannot, on any experimentalist approach to uncertainty, be fully tested from within the confines of any single iteration – any single judicial case or controversy. And society and the problems we delegate to agencies are too complex, too broad, and too politically polarizing to expect reasons that justify irrespective of predisposition or perspective. Until the judiciary adapts to this context, agency judgment and justification will remain obscure no matter how hard the look.

NOTES

3. 5 U.S.C. § 702 (creating a right of review in any “person suffering legal wrong because of agency action”).
5. 5 U.S.C. § 706(2). A total of six assumes that each subsection of 706(2) provides only one ground of reversal – an assumption that is certainly open to doubt. For example, an agency “action” could be “not in accordance with law” and yet be neither “arbitrary” nor “capricious.” But see id. at § 706(2)(A). I adopt this interpretation of Section 706 for argument’s sake.
7. Cf. *Bi-Metallic Invest. Co. v. State Bd. of Equalization*, 239 U.S. 442, 446 (1915) (holding that when an administrative authority takes action and a small number of persons are “exceptionally affected, in each case upon individual grounds,” the Due Process Clause requires that they be afforded some kind of adjudicatory hearing).

12. On the evasion of the APA’s mandate to reviewing courts that they “hold unlawful and set aside agency action, findings and conclusions found to” violate its strictures, see *Checkosky v. SEC*, 23 F.3d 452 (D.C. Cir. 1994) (remanding without vacating SEC action found to be “arbitrary and capricious”).
15. To be clear, the standard of proof in criminal matters – beyond a reasonable doubt – is supposed to be a higher standard of proof than is required in civil matters where the proof required is a “preponderance of the evidence.” See, e.g., *Jackson v. Virginia*, 443 U.S. 307, 317–319 (1979). How much higher and/or whether these two are best conceived in probabilistic terms is a question beyond our scope here, although it warrants mention that even common law courts obviously understood the spectral nature of proof.
19. Bose Corp., 466 U.S. at 500; Jackson, 443 U.S. at 324. Juries and trial judges have long been treated differently along this front by appellate courts’ standards of review, although it is less clear that that stems from their psychological makeup rather than their relative legal standing under Article III or the Sixth and Seventh Amendments. See *Weisgram v. Marley Co.*, 528 U.S. 440, 448 (2000); *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 376–378 (1996).
of deference to a trial court’s findings of fact in different language, requiring that they “must not be set aside [upon review] unless clearly erroneous” F.R.C.P. 52(a)(6).


26. Juries’ findings of fact were the first to be reviewed according to a “substantial evidence” standard before the standard was applied to administrative agencies. See, e.g., Gunning v. Cooley, 281 U.S. 90, 93 (1930). Eventually, the two standards were both described in similar terms, see Shields v. Utah Idaho Central R.R., 305 U.S. 177 (1938); NLRB v. Columbian Enameling & Stamping Co., 306 U.S. 292, 300 (1939), and then “substantial evidence” became something of a default standard courts would apply in reviewing agencies’ findings of fact. See Virginia Elect. & Power Co. v. NLRB, 319 U.S. 533 (1943); NLRB v. Waterman Steamship Corp., 309 U.S. 206, 224–226 (1940). Rule 52’s “clearly erroneous” standard replaced “substantial evidence” in the review of findings of fact by trial courts in 1937. While the two standards are nominally distinct, in practice they have proven remarkably similar. See Dickinson v. Zurko, 527 U.S. 150, 162–164 (1999). Both have been explicated in terms of the persons being reviewed. Cf. United States v. United States Gypsum Co., 333 U.S. 364, (1948) (holding that findings are “clearly erroneous” where the “reviewing court on the entire evidence is left with the definite and firm conviction that a mistake has been committed”); Universal Camera Corp. v. NLRB, 340 U.S. 474, 477 (1951) (explaining “substantial evidence” standard as “such relevant evidence as a reasonable mind might accept as adequate to support a conclusion”).


28. In Pacific States Box & Basket Co. v. White, 296 U.S. 176 (1935), Justice Brandeis for the Court held that a “presumption of validity” attached to agency action just as it did to legislation – at least for purposes of the Due Process clauses. Id. at 185–186. In National Labor Relations Bd. v. Universal Camera Corp., 340 U.S. 474 (1951), a rather pedantic opinion by Justice Frankfurter stated that no such presumption of validity could be reconciled with a reviewing court’s responsibilities under APA Section 10(e). Id. at 495–496.

29. Probabilism played little to no role in early discussions of substantial evidence, though, so the magnitude of this burden was never quite specified. Contrary evidence in the record was never tantamount to a lack of substantial evidence. See Denver Union Stockyard Co. v. United States, 304 U.S. 470 (1937). While some early cases did seem to involve the court in weighing the evidence itself, see, e.g., Federal Trade Comm’n v. Curtis Publishing Co., 260 U.S. 568 (1923); Federal Trade Comm’n v. Gratz, 253 U.S. 421 (1920), that notion quickly gave way to a more deferential “warrant in the record” model. See NLRB v. Hearst Pubs., Inc., 322 U.S. 111, 131 (1944).

30. The review of judgments of the Tax Court was long ago placed under a substantial evidence standard. See, e.g., Dobson v. Com’r, 320 U.S. 489 (1943).


32. See, e.g., Universal Camera Corp. v. NLRB, 340 U.S. 474 (1951). This may have been a selection effect, i.e., a function of which (biased) sample of matters were brought to court Allentown Mack Sales & Serv. v. NLRB, 322 U.S. 359 (1948).

34. See 5 U.S.C. §§ 554(a), 553(c).
35. See Ronald J. Allen, “Factual Ambiguity and A Theory of Evidence,” Northwestern University Law Review 88 (1994): 604, 616–630 (observing that the law of evidence and the system of proof burdens is best explained not in formal, probabilistic terms but rather as a way of eliciting and responding to the “best explanation” for the evidence introduced). Examples of this approach in arbitrariness review are plentiful. See, e.g., Mobil Oil Corp. v. FPC, 483 F.2d 1238, 1258–1263 (D. C. Cir. 1973); Corrosion Proof Fittings v. EPA, 947 F.2d 1201, 1213–1215 (5th Cir. 1991). Most especially, the legal tradition implies rights of confrontation – discovery, motions in limine, cross-examination, offers of proof, etc. Evidence accepted as such has been probed or at least occasioned the opportunity to do so.
36. See Fed. R. Evid. 401 (defining “relevant evidence” as “having any tendency to make the existence of any fact of consequence to the determination of the action more probable or less probable than it would be without the evidence”). Of course, the rules of evidence do not apply in agency proceedings. See Castillo-Villagra v. INS, 972 F.2d 1017, 1026 (9th Cir. 1992). Nevertheless, the tradition from which judicial review of agency action stems cannot but exert a significant influence on judges asked to assess a finding of fact, especially when the statutory standard explicitly invokes that tradition. Cf. id. at 1027–1029 (using the rules of evidence as a guide).
38. See ADAPSO, 745 F.2d at 683–684.
39. 745 F.2d at 684.
40. See 28 U.S.C. 2112(b) (defining the record as “the order sought to be reviewed or enforced, the findings or report upon which it is based, and the pleadings, evidence, and proceedings before the agency ... or such portions” as are filed pursuant to local rule, stipulation of the parties, or by motion of the court). Some courts provide for the filing of record excerpts along with the parties’ briefs. See Alan R. Gilbert, “Composition of Record on Review of Agency Action” under 28 U.S.C. § 2112(b) and Rule 16(a) of Federal Rules of Appellate Procedure, 32 A.L.R., (1977): 648, 649–650. Furthermore, the code provides that “[t]he agency ... may, at its option ... file in the court the entire record of the proceedings before it without abbreviation.” Id. at § 2112(c). This may well be, in the era of big data and digitization, a back-door for agencies to disguise or confuse the issues for review.
41. See Elizabeth Magill and Adrian Vermeule, “Allocating Power Within Agencies,” Yale Law Journal 120 (2011), pp. 1032, 1069 (arguing that the party who assembles the “record” in the first instance, i.e., the hearing neutral or “Administrative Law Judge,” is empowered by substantial evidence review).
Because most on-the-record proceedings will be governed by the APA, see 5 U.S.C. § 559, unless “otherwise provided by statute, the proponent of a rule or order has the burden of proof.” Id. at § 556(d). Thus, when substantial evidence review occurs, it will be of a proceeding and findings of fact where the agency—normally the proponent of its rule/order—will have had to carry a burden of proof/persuasion with “reliable, probative, and substantial evidence.” Id.


See, e.g., Greater Boston, 444 F.2d at 851 (calling the court’s review a “supervisory function”); Environmental Defense Fund, Inc. v. Ruckelshaus, 439 F.2d 584, 597–598 (D.C. Cir. 1971); Kennecott Copper Corp. v. EPA, 462 F.2d 846, 850–851 (D.C. Cir. 1972); International Harvester Co. v. Ruckelshaus, 478 F.2d 615, 641–647 (D.C. Cir. 1973); Ethyl Corp. v. EPA, 541 F.2d 1, 96 (D.C. Cir. 1976) (Wilkey, J., dissenting).


Greater Boston, 444 F.2d at 841, 851.


57. In Citizens to Preserve Overton Park v. Volpe, 401 U.S. 402, the Court held that § 706(2)(A)’s review for arbitrariness presumed that agencies would substantiate their own reasons and decision-making. Id. at 419–420. Agencies like EPA pioneered the creation of records in “informal” proceedings consistent with this expectation. See, e.g.,


60. See, e.g., Friends of the Earth v. AEC, 485 F.2d 1031, 1032–1033 (D.C. Cir. 1972); International Harvester Co. v. Ruckelshaus, 478 F.2d 615, 650–653 (D.C. Cir. 1973) (Bazelon, J., concurring); Ethyl Corp. v. EPA, 541 F.2d 1, 66–68 (D.C. Cir. 1976) (Bazelon, J., concurring).


62. See Motor Vehicle Mfrs. Ass’n v. State Farm Automobile Ins. Co., 463 U.S. 29, 43 (1983). The Court was also explicit that it did “not view as equivalent the presumption of constitutionality afforded legislation drafted by Congress and the presumption of regularity afforded an agency in fulfilling its statutory mandate.” 463 U.S. at 43 n.9. The implication was that the agencies’ was the weaker presumption.

63. State Farm, 463 U.S. at 43. Quoting one of the original hard look precedents, the Court further observed that “[a]n agency’s view of what is in the public interest may change, either with or without a change in circumstances. But an agency changing its course must supply a reasoned analysis.” Id. at 57 (quoting Greater Boston, 444 F.2d at 852). Some, therefore, referred to State Farm’s formula as the “swerve” doctrine.

64. See Edley, supra, at 61–66.

65. State Farm, 463 U.S. at 43–44.

66. See State Farm, 463 U.S. 29, 43–44 (emphasis added).

67. See State Farm, 463 U.S. at 46–57 (reviewing arguments and information from the fifteen year history of Standard 208’s proposal finalization, repeal, subsequent re-proposals and re-finalization).

68. The further irony is that the Safety Act nowhere explicitly levied the substantial evidence requirement the State Farm Court assumed. See State Farm Mutual Auto Ins. v. Dept. of Transp., 680 F.2d 206, 218–220 (D.C. Cir. 1982) (noting the Act’s various interpretations by reviewing courts and concluding that there was no practical difference because arbitrariness and substantial evidence review are substantially the same inquiry). Imposing the substantial evidence scope of review, though, arguably ignored the Court’s own admonitions in Vermont Yankee and Florida East Coast Ry.


70. “The standard [being] rescinded was the subject of approximately 60 notices of proposed rulemaking, hearings, amendments, and the like between 1969 and 1981.” State Farm, 680 F.2d at 209. Indeed, throughout the Safety Act’s standard-setting processes, the statute’s record requirement became of singular importance. See Automotive Parts & Accessories Ass’n v. Boyd, 407 F.2d 330, 336–338 (D.C. Cir. 1968); Chrysler Corp. v. Dept. of Transp., 472 F.2d 659, 669, 681 n.29 (6th Cir. 1972); Mashaw and Harfst, supra, at 156–171.

71. Expertise is nothing more (or less) to the legal system than “knowledge, skill, experience, training or education” that “will assist the trier of fact to understand the evidence or to determine a fact in issue.” Fed. R. Evid. 702. Standard 208’s accumulated record of projected vehicle price increases, avoided fatalities, probable forcing of
technologies, etc., was deficient not because it lacked expert testimony but rather because it lacked expert testimony answering the exact questions framed by State Farm’s lawyers. See Mashaw and Harfst, supra, at 210–223. Lower federal courts have routinely found themselves on this slippery slope when reviewing agency action that relies on expertise. See American Radio Relay League, Inc. v. FCC, 524 F.3d 227, 245–248 (D.C. Cir. 2008) (Kavanaugh, J., concurring and dissenting in part).

72. See, e.g., American Radio Relay League, Inc. v. FCC, 524 F.3d 227, 237 (D.C. Cir. 2008) (“It would appear to be a fairly obvious proposition that studies upon which an agency relies in promulgating a rule must be made available during the rulemaking in order to afford interested persons meaningful notice and an opportunity for comment.”); United States v. Nova Scotia Food Prods. Corp., 568 F.2d 240, 251 (2d Cir. 1977) (concluding that “when the pertinent research material is readily available and the agency has no special expertise on the precise parameters involved” there is no “reason to conceal the scientific data relied upon from the interested parties”).

73. In a recent rulemaking adjusting the applicability of the Clean Air Act’s “new source review” requirements to emitters of the six principal greenhouse gases, EPA received over 400,000 comments in response to its proposed rule. See U.S. EPA, Final Rule: Prevention of Significant Deterioration and Title V Greenhouse Gas Tailoring Rule, 75 Fed. Reg. 31514 (2010).


75. Cf. Edley, supra, at 56 (“[N]o natural law of management assures us that a collection of experts will produce worthy administration. The whole is neither more nor less than the sum of its parts. It is an altogether different phenomenon.”).


79. See 42 U.S.C. § 7602(g) (defining “air pollutant” to include “any physical, chemical ... substance or matter which is emitted into or otherwise enters the ambient air” as an “air pollution agent”). No definition of air pollution agency appears in the act.

80. Massachusetts, 549 U.S. at 507–509.

81. The “record” was the petition itself, a Federal Register notice requesting comment on the petition, the public comments thereon received, and the Federal Register notice denying the petition. See U.S. EPA, Control of Emissions From New Highway Vehicles and Engines: Notice of Denial of Petition for Rulemaking, 68 Fed. Reg. 52922 (2003). EPA was under no obligation to confine its explanation for the denial to this “record,” though.

82. 68 Fed. Reg. at 52927. The dissent made much of EPA’s discussion of the lingering doubts about climate change’s causation. See Massachusetts, 549 U.S. at 553–555. It is, however, a mistake to infer that EPA had argued, even in 2003, that climate change was probably not in progress. What the agency argued was that causality,
especially as between the several different agents, could not be “unequivocally established.” Id. at 553. That would be true of any causal relationship, though.

83. Two prior EPA General Counsels had taken the position that the Clean Air Act did authorize EPA to regulate greenhouse gases as air pollutants. 68 Fed. Reg. at 52925. However, the agency’s amended interpretation was explained not on the basis of changed personnel, but rather in light of a then-recent Supreme Court precedent, FDA v. Brown and Williamson Tobacco Corp., 529 U.S. 120 (2000).


85. Had that been the agency’s “finding,” one would have expected to see more attention to the GHGs not even mentioned by the petitioners – emissions like water vapor, black carbon, sulfur hexafluoride (SF6), and perfluorocarbons (PFCs).

86. The Act even explicitly stated that the Administrator had to make a “judgment” whether the cited “air pollutants” either “cause” or “contribute to” public endangerment. See 42 U.S.C. § 7521(a)(1); Massachusetts, 549 U.S. at 549 (Scalia, J., dissenting).


88. “[T]he statute says nothing at all about the reasons for which the Administrative may defer making a judgment – the permissible reasons for deciding not to grapple with the issue at the present time.” Massachusetts, 549 U.S. at 552 (Scalia, J., dissenting).

89. See Massachusetts, 549 U.S. at 535.

90. See 68 Fed. Reg. at 52925, 52931; Massachusetts, 549 U.S. at 533 (“[O]nce EPA has responded to a petition for rulemaking, its reasons for action or inaction must conform to the authorizing statute.”).


93. Cf. Hart and Honoré, supra, at 264 (“[T]here is a logical absurdity in asking whether the risk of further harm, arising from a harmful situation which a reasonable man would not have created, would itself have deterred a reasonable man from acting.”).

94. Massachusetts, 549 U.S. at 533 (“EPA has refused to comply with th[e] clear statutory command.”).


96. See Ethyl Corp. v. EPA, 541 F.2d 1 (D.C. Cir. 1976); Lead Industries Ass’n, Inc. v. EPA, 647 F.2d 1130, 1156–1167 (D.C. Cir. 1980); American Petroleum Inst. v. Costle, 665 F.2d 1176, 1184–1187 (D.C. Cir. 1981). In Ethyl Corp., the D.C. Circuit confronted lead as a fuel additive and its regulation under Title II of the Clean Air Act. The argument was made in contesting EPA’s endangerment finding there that lead entered the body from so many different sources that regulating it as a fuel additive was pointless. Id. at 8–9. It was only with the evidence gathered suggesting that airborne lead from fuel additives was a significant source of human blood lead contamination that the en banc majority upheld EPA’s determination. See id. at 31–32. Likewise, in the famous Benzene case, Justice Stevens himself (the author of the majority opinion in
Massachusetts) wrote for a plurality considering an ambiguous mandate in the Occupational Safety and Health Act that the agency protect workers from risk that the agency must first find that exposure to toxics like benzene create a “significant” risk and not simply a trivial or redundant risk. See Industrial Union Dept., AFL-CIO v. American Petroleum Inst., 448 U.S. 607, 639–47 (1980) (Stevens, J., joined by Burger, Stewart, Powell, JJ.).


100. Massachusetts, 549 U.S. at 552-53 (Scalia, J., dissenting) (citing Chevron U.S.A., Inc. v. Natural Resources Defense Council, Inc., 467 U.S. 837 (1984)). The applicability of the Chevron doctrine to this particular agency interpretation need not concern us here as there were (and are) other deference doctrines that were equally appropriate under the circumstances.


103. See, e.g., Lead Industries Ass’n v. EPA, (D.C. Cir. 1980).

104. See, e.g., Whitman v. American Trucking Ass’ns, Inc., 531 U.S. 457 (2001) (affirming that EPA cannot consider cost in the setting of National Ambient Air Quality Standards under Clean Air Act § 109(a) consistent with long standing precedent to that effect).


107. I use “epistemic” here as an umbrella for the set of all risks of error that a decision-maker confronts which is the result of its own competence or cognition.

108. Cf. Scott Brewer, “Scientific Expert Testimony and Intellectual Due Process,” Yale Law Journal 107 (1998), p. 1535 (arguing that wherever arguments are close, there are compelling reasons to doubt that nonexpert legal decision-makers can acquire information from an expert in any way that is both legally and epistemically justified); Christopher Tarver Robertson, “Blind Expertise,” New York University Law Review 85 (2010), pp. 174, 181–201 (arguing that party-selected experts are often unconsciously biased and that trial judges lack the necessary competence to sort out the testimony that is accurate and truth-tracking from that which is erroneous or error-prone).


111. Reviewing courts have long insisted that an agency’s “post hoc rationalizations,” i.e., explanations or evidence occasioned by the litigation itself, are not to be trusted. See, e.g., SEC v. Chenery Corp., 332 U.S. 194, 197–199 (1947); American Textile Mfrs., Inc. v. Donovan, 452 U.S. 490, 539 (1981).

112. For the more general claim that combining too much of a cognitive load into one decisional point is a failure of institutional design, see Frederick Schauer, “Do Cases Make Bad Law?” University of Chicago Law Review 73 (2006), p. 883.


117. See Vermeule, supra, at 118–48 (arguing that the “common conceptual mistake” in accounts of judicial review is “the undefended assumption that sustained judicial coordination on a particular interpretive approach is feasible” and that that assumption is, in fact, unwarranted given the heterogeneity and independence of the judiciary).


119. See Eric Talley, “Precedential Cascades: An Appraisal,” *Southern California Law Review* 73 (1999), pp. 87, 99–120. “If judges have a hard time avoiding what they see as the right result for the particular case in all of its contextual richness, and if they are at the same time making law for future cases, then the combination of the salience of the particular case and the pull to decide it correctly may produce a rule that is unrepresentative of the full range of future cases that can be expected to be decided under it.” Schauer, supra, at 900.

120. See, e.g., John H. Langbein, “The German Advantage in Civil Procedure,” *University of Chicago Law Review* 52 (1985), pp. 823, 833 n.31 (“Wigmore’s celebrated panegyric – that cross-examination is ‘the greatest legal engine ever invented for the discovery of truth’ – is nothing more than an article of faith.... In the hands of many of its practitioners, cross-examination is not only frequently truth-defeating or ineffectual, it is also tedious, repetitive, time-wasting, and insulting.”).


122. See, e.g., Philip Kitcher, *Abusing Science: The Case Against Creationism* (Cambridge, Mass.: MIT Press, 1982); cf. Brewer, supra, at 1590–1596 (observing the compound challenge for non-experts in judging the competence of experts in that expertise comes in degrees and domains and that it is constantly changing through updates, overthrow, etc.).


125. See Stephen Breyer, “Judicial Review of Questions of Law and Policy,” *Administrative Law Review* 38 (1986), p. 363, 380. Seven years before *State Farm*, the D.C. Circuit confronted this paradox of hard look review. A majority of the court sitting en banc answered that it was a reviewing court’s duty, in the face of its own incompetence on matters of scientific evidence, to “look at the decision not as the chemist, biologist, or statistician,” but rather to educate itself with the record and a “conscientious awareness” of the “limited nature” of its “narrowly defined duty of holding agencies to certain minimal standards of rationality.” *Ethyl Corp v. EPA*, 541 F.2d 1, 36 (D.C. Cir. 1976). While this sounds judicious, it masks the underlying epistemic paradox of hard look review with what amounts to euphony.


130. See, e.g., *Bennett v. Spear*, 520 U.S. 154 (1997). This was true before the APA and regardless of the agency’s labeling of its norm. See, e.g., *Columbia Broadcasting System, Inc. v. United States* 316 U.S. 407, 415–420 (1942) (holding that the FCC’s promulgation of its “Chain Broadcasting Regulations” by “order” was reviewable in an equity suit by those who would have to seek re-licensing subject to those regulations).

131. 5 U.S.C. § 551(4). Typical in this regard are agency policy statements to the public that, in practical effect, tie the agency to some particular judgment or course of conduct. See, e.g., *Community Nutrition Inst. v. FDA*, 818 F.3d 943, 948–950 (D.C. Cir. 1987) (enforcement guideline); *Alaska Professional Hunters Ass’n v. FAA*, 177 F.3d 1030 (D.C. Cir. 1999) (long-standing interpretation of agency regulation).


137. Dorf and Sabel, supra, at 395.


141. See, e.g., *Starr v. FAA*, 589 F.2d 307 (7th Cir. 1978); *Rombough v. FAA*, 594 F.2d 893 (2d Cir. 1979); *Gray v. FAA*, 594 F.2d 793 (10th Cir. 1979); *Keating v. FAA*, 610 F.2d 611 (9th Cir. 1979).

142. See *Aman v. FAA*, 856 F.2d 946, 954 n.7 (7th Cir. 1988). While other pilots became protected from such “age-related employment discrimination” through the federal Age Discrimination in Employment Act, commercial airline pilots and their employers remained governed by the FAA’s Age 60 rule. See *Aman v. FAA*, 856 F.2d 946, 954–956 & n.13 (7th Cir. 1988).

143. See 24 Fed. Reg. at 9772 (“[I]ncapacity ... cannot be predicted accurately as to any specific individual on the basis of presently available scientific tests and criteria. On the contrary, the evidences of the aging process are so varied in different individuals that it is not possible to determine accurately with respect to any individual whether the presence or absence of any specific defect in itself either led to or precluded a sudden incapacitating attack.”)

144. Every time the agency explained the rule’s original justification the reviewing court accepted it as valid and sufficient. See, e.g., *Professional Pilots Fed’n v. FAA*, 118 F.3d 758, 765 (D.C. Cir. 1997)(“There are no tests ... that can accurately determine the risk of an apparently healthy but older pilot suddenly being stricken by any one of the many potentially disabling conditions that may accompany advancing age.”). In the final iteration, the Professional Pilots’ Association proposed an evaluative tool—a means of screening blood for early indications of stroke or heart failure—but FAA’s rejection of the tool as incompletely developed was upheld over the pilots’ challenge. See *Yetman v. Garvey*, 261 F.3d 664, 673–675 (7th Cir. 2001).

145. The Federal Aviation Act arguably levied a “substantial evidence” requirement in the review of petition denials. See *Aman v. FAA*, 856 F.2d 946, 949–951 (7th Cir. 1988). The courts interpreted substantial evidence review in petition denial cases to require that some but not necessarily the weight of the evidence support FAA. See id. at 951–954 and n.6.


147. After decades of successive arbitrariness challenges, one court of appeals seemed to confess the agency’s “power to establish a rigid policy” and to adhere to it “inflexibly” so long as its challengers could not prove FAA had given their claims no consideration at all. See *Yetman*, 261 F.3d at 679. That is probably more agency autonomy than we have reason to prefer.


149. In challenging the rational basis for any new rule, the pilots would presumably be in better position to draw out the FAA’s alleged inconsistencies in
allowing younger pilots who had suffered a first heart attack, and younger pilots that suffered from alcoholism to fly commercially in the U.S. as long as they had obtained the right medical assurances while at the same time grounding pilots aged 60 + 1 day with exemplary health and safety records. See Yetman, 261 F.3d at 671–672.

150. Baker v. FAA, 917 F.2d 318, 322 (7th Cir. 1990).

151. In a GAO report to Congress it was revealed that by the end of 1989, 418 pilots had petitioned in 67 different proceedings seeking exemptions from the rule. See U.S. General Accounting Office, Aviation Safety: Information on FAA’s Age 60 Rule for Pilots 2 (GAO/RCED-90-45FS) (Nov. 1989).


155. While courts had been reviewing petition denials for an “abuse of discretion,” see, e.g., Starr v. FAA, 589 F.2d 307 (7th Cir. 1978), that standard was jettisoned once FAA started keeping records conducive to “substantial evidence” review – even though the agency arguably was not bound to do so. See Aman v. FAA, 856 F.2d 946, 951 (7th Cir. 1988).

156. See Pub. L. No. 110–135, 121 Stat. 1450 (2007). In a bizarre twist, the statute provided no “retroactivity” to protect pilots who might miss its amnesty by mere days. See Jones v. Air Line Pilots Ass’n, Int’l, et al., 642 F.3d 1100, 1102 (D.C. Cir. 2011).


158. See Frank H. Knight, Risk, Uncertainty, and Profit (Boston: Hart, Schaffner and Marx, 1921), pp. 19–21, 197–232.


160. But see Yetman v. Garvey, 261 F.3d 664, 678 (7th Cir. 2001) (“If the FAA was justified in imposing the Age Sixty Rule in the first place, then we cannot say that simply because it is the rule itself which blocks the generation of data necessary to consider the propriety of granting exemptions to the rule, that it was unreasonable for the FAA to find that it lacks that data.”).

161. In her dissent in Professional Pilots Federation, Judge Wald highlighted the point made in text and argued further that an experiment with some carefully screened group of 60-plus pilots could have improved agency knowledge. See Professional Pilots Fed’n v. FAA, 118 F.3d 758, 775 (D.C. Cir. 1997) (Wald, J., dissenting).


