THE PARLIAMENT OF THE EXPERTS

ADRIAN VERMEULE†

ABSTRACT

In the administrative state, how should expert opinions be aggregated and used? If a panel of experts is unanimous on a question of fact, causation, or prediction, can an administrative agency rationally disagree, and on what grounds? If experts are split into a majority view and a minority view, must the agency follow the majority? Should reviewing courts limit agency discretion to select among the conflicting views of experts, or to depart from expert consensus? I argue that voting by expert panels is likely, on average, to be epistemically superior to the substantive judgment of agency heads, in determining questions of fact, causation, or prediction. Nose counting of expert panels should generally be an acceptable basis for decision under the arbitrary and capricious or substantial evidence tests. Moreover, agencies should be obliged to follow the (super)majority view of an expert panel, even if the agency's own judgment is to the contrary, unless the agency can give an epistemically valid second-order reason for rejecting the panel majority's view.

TABLE OF CONTENTS

---

† John H. Watson Professor of Law, Harvard Law School. Presented at the *Duke Law Journal*'s 2009 Administrative Law Symposium. Thanks to Lisa Bressman, Glenn Cohen, Richard Fallon, Lisa Heinzerling, Matthew Stephenson, Jonathan Weiner, and the conference participants for helpful comments, and to Colleen Roh for excellent research assistance.

INTRODUCTION

In the administrative state, a great deal of agency decisionmaking draws upon the aggregate view of a group of experts, especially when there is an expert "consensus." Under what conditions is this practice sensible, or not? If experts are unanimous on a complex question of fact, causation, or prediction, can an agency rationally disagree, and on what grounds? If experts are split into a majority view and a minority view, must the agency follow the majority? Should reviewing courts limit agency discretion to select among the conflicting views of experts, or to depart from expert consensus?

To come to grips with these problems, I will focus on advisory panels of scientific and technical experts and their role in administrative law. How should the views and votes of the members of these expert panels, whether unanimous or conflicting, be aggregated and incorporated into agency decisionmaking and judicial

review of agency action? The problem is ubiquitous in administrative law; it also underlies several major episodes that, observers suggest, illustrate the politicization of science under the Bush administration.

To structure the discussion, I will examine problems like the following:

• In 1979, the Environmental Protection Agency (EPA) promulgated ozone standards that had been approved by a bare majority, six of eleven members, of a scientific advisory panel;[1] similarly, in 1997, EPA revised its standards for ozone and particulate matter and justified the revisions in part by reference to the consensus or majority views of the panel.[2] In 2006, however, in issuing a final rule on particulate matter, the Administrator declined to follow the recommendations of the same panel, in part because the panel was not unanimous; two out of the twenty-two members had dissented.[3] Are these positions consistent? If they are not, which was correct?

• In the second case, when twenty of twenty-two scientists on the advisory panel recommended action that EPA declined to take, would EPA's decision look better or worse if the panel had been unanimous?

• A statute gives benefits to miners who suffer from pneumoconiosis, or black lung disease. In a black lung case, three experts give a diagnosis that there is no pneumoconiosis, whereas two experts disagree.[4] May the administrative law judge say that the three diagnoses outweigh the two, and award benefits on that basis? May the judge say that the two diagnoses outweigh the three?

• The National Marine Fisheries service must decide whether to list a genetically significant subpopulation of salmon as threatened. It convenes a panel of sixteen scientists to decide, among other things, whether certain technical measures will avert the danger. The panel splits about evenly.[5] Can the agency adopt either view?

Administrative law has no consistent view about how expertise should be aggregated in cases like these. In particular, judicial decisions are inconsistent about whether agencies may or must count

1. Am. Petroleum Inst. v. Costle, 665 F.2d 1176, 1182 (D.C. Cir. 1981).

2. Am. Trucking Ass'ns v. EPA, 283 F.3d 355, 367, 376–79 (D.C. Cir. 2002).

3. National Ambient Air Quality Standards for Particulate Matter, 71 Fed. Reg. 61,144, 61,174 n.44 (Oct. 17, 2006) (codified at 40 C.F.R. § 50.6 (2008)).

4. Stalcup v. Peabody Coal Co., 477 F.3d 482, 484 (7th Cir. 2007).

5. Or. Natural Res. Council v. Daley, 6 F. Supp. 2d 1139, 1148 (D. Or. 1998).

noses—whether agencies may or must aggregate expertise by treating expert views as votes, and adopting the view for which a majority or supermajority of experts casts their votes. My major normative claim is that nose counting by agencies is usually permissible and should sometimes be mandatory; I mean to praise the parliament of the experts. In some cases judges say that agencies may not count noses, but I will say that agencies should be allowed to do so. In other cases, I will even claim that agencies should be presumptively obligated to adopt the views of a majority or supermajority of the experts on the agency's advisory panels, absent the right kind of epistemic reason for departing from those views.

In short, administrative law sometimes, although inconsistently, has a horror of a quantitative approach to the incorporation of expertise, and a strong preference for a qualitative, reason-based approach. But I will claim that when agencies are uncertain of facts, causation, or future consequences of alternative policies, following the consensus or majority view of experts is a perfectly rational decisionmaking strategy. Indeed, nose counting of experts is presumptively superior to fact-finding based on the agency's own substantive reasons or views, and it should prevail unless there is a sound epistemic reason to believe that the agency is better positioned than the expert panel to determine the relevant facts.

Much of the confusion and inconsistency surrounding these issues is dispelled by a simple distinction between first-order reasons and second-order reasons. First-order reasons are reasons for thinking that a particular claim of fact, causation, or prediction is correct. Second-order reasons are reasons for thinking that the epistemic capacities of (a group of) experts are such that their first-order views are more or less likely to be correct, compared to the agency's first-order views. When courts say that agencies who have counted expert noses have given no valid reason for the decision, the courts overlook that nose counting, although it offers no first-order reason, is a perfectly rational second-order decisionmaking strategy for agencies that lack first-order competence. Furthermore, when the weight of expert opinion on matters of fact, causation, or prediction tilts in one direction, agencies should be required either to follow the expert opinion or else give a valid second-order reason to think that expert opinion is epistemically suspect. Contrary to much of the

current law, I suggest that the agency's own first-order reasons are not an adequate basis for departing from the expert consensus.[6]

Part I lays out some preliminary assumptions and offers a synopsis of relevant administrative law doctrine. Part II asks when agencies *may* count noses, focusing on cases in which laws bar agencies from counting expert noses. I suggest that nose counting, under identifiable conditions, is a perfectly sound second-order epistemic strategy; when those conditions are met, courts should hold that nose counting is a rational basis for agency decisionmaking. Part III describes the conditions under which agency nose counting is inadequate. Part IV asks when agencies *must* count noses, focusing on cases in which agencies attempt to depart from the findings reached by a consensus or a majority of experts. I suggest an easily implemented presumption: agencies should not be permitted to depart from the findings of expert panels unless they can give a valid second-order reason to think that the consensus or majority view of experts as to matters of fact is not epistemically reliable. A brief conclusion follows.

I. ASSUMPTIONS AND THE LAW

A. Preliminaries

Critics of the Bush administration's decisionmaking decry the "politicization" of science.[7] The anecdotes are many, and unsettling, but the critics are rarely clear or explicit about the theory that underlies the critique, and the very variety of the anecdotes underscores that the problems are highly heterogeneous. In many cases,[8] expert consensus is used as an implicit benchmark for

6. *Cf.* E. Donald Elliott, *Strengthening Science's Voice at EPA*, 66 LAW & CONTEMP. PROBS. 45, 51 (Autumn 2003) ("When the scientists at EPA, such as the Science Advisory Board, have refused to approve the Agency's scientific rationale, a court should consider that refusal in giving lesser deference to the agency's decision."). Although I agree with the broad thrust of Elliott's suggestion, I hope to make the idea of "lesser deference" more precise, in part by distinguishing between the agency's first-order reasons (which will not be a valid basis for departing from the expert panel's recommendations) and the agency's second-order reasons (which may be a valid basis for doing so, under some identifiable conditions).

7. *See, e.g.*, Sidney A. Shapiro, *OMB and the Politicization of Risk Assessment*, 37 ENVTL. L. 1083, 1084 (2007) (discussing "[t]he [Bush] administration's efforts to politicize science").

8. Shapiro offers two examples of "science denial" and "politicization" by the Bush administration. Shapiro, *supra* note 7, at 1086–87. The first is that the administration "refus[ed] to acknowledge or act on the overwhelming scientific evidence of global climate change . . . ." *Id.* at 1086. In the second, "[t]he Food and Drug Administration . . . refused to approve the

determining that science has been politicized, yet that benchmark is undertheorized.

Somewhat paradoxically, I believe that the best response to this heterogeneity is to focus on a crucial subset of the relevant issues. If the subset is well chosen, it will illuminate the structure of the broader problem while avoiding the morass of detail in which these critiques sometimes sink. Accordingly, I focus on the relationship between agencies and the scientific and technical panels that advise them, and on judicial review of agency action involving decisions that follow or reject the recommendation of an expert panel.[9]

Some preliminary assumptions are necessary to clarify the questions. I will assume that agency policymaking is a two-stage process, encompassing both fact-finding and evaluation. At the first stage the agency finds "facts"—namely, discrete adjudicative facts, background conditions or legislative facts, causal theories, and predictions about the consequences of alternative policies. Expertise, in this framework, just means a higher probability of getting the facts right.

Once the facts are determined, agencies decide what to do given those facts. At this second stage, agencies apply some decision rule determined either by their own preferences or by the preferences of Congress and the president, in some mix. Expertise has no relevance at this second stage, which involves the application to the facts of values or preferences as to which there may be irreducible conflict or disagreement. This is, of course, a hopelessly simplified picture of the policymaking process. However, the simplification is useful for my purposes, which are legal and normative rather than conceptual. I do not aim to give a philosophically adequate account of the elements of public choice, but rather to make sense of the relationship between agencies and expert panels in the administrative state, under extant statutes. As we will see, this distinction between expert fact-finding and agency evaluation best reconciles some major features of the relevant statutes and of the broader legal landscape.

---

emergency contraceptive Plan B, despite the fact that two scientific advisory committees had overwhelmingly found that the drug was safe and effective." *Id.* at 1086–87.

9. Given this comparison between expert panels and agencies, I put aside other mechanisms for decisionmaking that might prove superior to both, such as the use of regression-based algorithms. For an explanation, and evidence that such algorithms outperform experts even in tasks saturated with uncertainty, see PHILIP E. TETLOCK, EXPERT POLITICAL JUDGMENT 77 (2005).

Within this two-stage framework, I focus solely on the first stage—the determination of facts. Nothing in the following discussion suggests that agencies must defer to expert panels in evaluating the outcomes of possible policies. An example involves the recurrent question of how agencies should set their regulatory priorities. Reviewing courts have afforded agencies broad discretion to allocate resources and pick their targets, even where expert advisory panels have recommended a course of action.[10] To the extent that agency priority-setting and, more generally, agency evaluation of alternative policies are premised on factual or causal claims or predictions as to which expert panels have been asked to make findings, the analysis applies in full.

A corollary is that I will bracket the question of how much political accountability over agency decisionmaking is desirable. For any desired level of political accountability, and for any desired distribution of policymaking authority among Congress, the president, and agencies, it is better, from the social point of view, to get facts right than to get them wrong. Congressional and presidential preferences may influence the choice of policies, given certain facts, but they should not influence the factual component of agency decisionmaking. There is no social benefit, and real social cost, when an agency claims that a species is numerous when it is actually extinct, or claims that the science surrounding climate change is uncertain when it is not. If Congress and the president do not think that species loss or climate change are problems worth addressing, agencies may so decide, subject to the constraints of existing law. But agencies may not disguise their policy preferences in the language of fact, a course of action that hampers political accountability by making it more difficult for legislative and executive principals to monitor the agencies.[11] In the framework I suggest, by contrast, agencies would have incentives to openly explain their normative differences with

10. *See, e.g.*, Int'l Union, UAW v. Chao, 361 F.3d 249, 255 (3d Cir. 2004) (upholding an Occupational Safety and Health Administration decision not to make rules to regulate occupational exposure to machining fluids, despite an advisory committee's contrary recommendation, on the ground that the agency had rationally set other priorities). On the general issue of judicial deference to agency priority-setting, see Massachusetts v. EPA, 549 U.S. 497, 533–34 (2007); Heckler v. Chaney, 470 U.S. 821, 831–32 (1985); Eric Biber, *The Importance of Resource Allocation in Administrative Law*, 60 ADMIN. L. REV. 1 *passim* (2008).

11. Allentown Mack Sales & Serv., Inc. v. NLRB, 522 U.S. 359, 376 (1998) ("An agency should not be able to impede judicial review, and indeed even political oversight, by disguising its policymaking as factfinding.").

expert panels, producing a kind of normative transparency that reduces the costs of monitoring their commitments and behavior.

Of course there is a blurry line between evaluation on the one hand and questions of fact on the other;[12] the two components of agency policymaking lie on a continuum and are often intertwined. When agencies set rates for regulated utilities, the rate that will give a reasonable return is a question of policy sitting atop several questions of fact. When agencies choose uncertainty parameters for estimates of serious risks, the choice of parameter incorporates complex judgments about the risk and harms of error in one direction or another.[13] Presumptions and burdens of proof and persuasion, as to matters of fact, can themselves rest on agency judgments of policy. Agencies sometimes blur the categories deliberately, engaging in a "science charade"[14] that disguises evaluation as expert fact-finding.

Nonetheless, there are good pragmatic reasons for using this philosophically dubious distinction. The category of questions of fact is written into the text[15] and structure of the Administrative Procedure Act (APA), and it pervasively shapes the doctrine of administrative law, so there is little choice but to make the best of it. In many cases, one can untangle value choices from factual, causal, and predictive questions, with some work, and there are also many clear cases of factual, causal, and predictive judgments. Whether a black lung claimant has or will develop pneumoconiosis is predominantly a question of fact, causation, or prediction, in any ordinary sense of those terms. So too with the question whether reducing the permissible level of particulate matter in the air will result in less asthma. The line between fact and evaluation is sometimes difficult to draw, sometimes not; the difficulty of the line-drawing exercise is not unique to this area, but it is a challenge across all of administrative law.

---

12. *See generally* HILARY PUTNAM, THE COLLAPSE OF THE FACT/VALUE DICHOTOMY AND OTHER ESSAYS (2002).

13. Thomas O. McGarity, *Substantive and Procedural Discretion in Administrative Resolution of Science Policy Questions: Regulating Carcinogens in EPA and OSHA*, 67 GEO. L.J. 729, 748–49 (1979).

14. THOMAS O. MCGARITY & WENDY E. WAGNER, BENDING SCIENCE: HOW SPECIAL INTERESTS CORRUPT PUBLIC HEALTH RESEARCH 269 (2008) (explaining that policymakers are "usually quite reluctant" to abandon the "science charade"); Wendy E. Wagner, *The Science Charade in Toxic Risk Regulation*, 95 COLUM. L. REV. 1613, 1617 (1995) (describing the "'science charade,' in which agencies exaggerate scientific contributions in setting toxic standards to avoid accountability for underlying policy decisions").

15. 5 U.S.C. § 706(2)(F) (2006).

A common mechanism by which agencies determine facts, in cases of any complexity, is to appeal to expert opinion. Informally, the agency may cite studies or simply make claims about the views of experts in the field. More formally, the agency may convene an expert panel to offer views. In the most important cases, agencies are obligated to do so by statute. When revising National Ambient Air Quality Standards (NAAQS), EPA must consider the report of its permanent Clean Air Scientific Advisory Committee (CASAC) and explain any departure from the committee's findings and recommendations.[16] Likewise, Congress established an Advisory Commission on Childhood Vaccines (ACCV) within the Department of Health and Human Services (HHS). If the secretary of HHS receives a recommendation from the Commission, the secretary must either conduct a rulemaking in accordance with the recommendation or publish a "statement of reasons" for refusing to do so in the Federal Register.[17]

The latter class of cases, for which Congress itself has set up panels by statute or set constraints on the panels agencies may choose, is my central focus. I will not address how expert panels should ideally be designed, what rules they should follow, or how their memberships should be chosen. Although these questions of statutory reform and institutional design are critical,[18] it is also important to ask how administrative law and judicial review of agency action should be structured *given* the panels that Congress and the agencies have actually set up. As we will see, the structure, composition, and procedures that panels use will themselves have direct implications for administrative law and judicial review.

B. How the Law Stands

Under administrative law, how does expert consensus or disagreement affect agency decisionmaking? Under the APA, several legal standards can potentially become relevant in these cases, with a great deal of overlap among them.

In agency proceedings, the default standard of proof is preponderance of the evidence,[19] unless statutes specify otherwise. I

16. 42 U.S.C. § 7607(d)(3) (2006).

17. *Id.* § 300aa-14(c)(2).

18. For an excellent treatment of statutory reform and institutional design, see MCGARITY & WAGNER, *supra* note 14, at 262–75.

19. *See, e.g.*, Steadman v. SEC, 450 U.S. 91, 102 (1981) (interpreting 5 U.S.C. § 556(d)).

will assume throughout that this is so, ignoring rare cases in which statutes or constitutional rules mandate that agencies decide by clear and convincing evidence. Once the agency has made its findings, the APA instructs courts to set aside agency decisionmaking that is "arbitrary, capricious, [or] an abuse of discretion."[20] Agencies must always, at a minimum, offer a reasoned basis for their factual findings and policy choices. Where proceedings are on the record, however, a reviewing court must find substantial evidence for the agency's factual findings.[21] The traditional understanding of the substantial evidence standard is that the agency prevails if it offers "such relevant evidence as a reasonable mind might accept as adequate to support a conclusion,"[22] analogous to the standard of review courts use to review jury findings.[23]

It is not at all clear how much, if at all, the substantial evidence standard of review for proceedings on the record ultimately differs from the general arbitrary or capricious standard. Although refined distinctions between these two standards have been drawn, the best view, and a widespread one, is that the substantial evidence standard is just a more specific application of the general arbitrary and capricious standard;[24] in practice, the two tend to collapse into one another.[25] Although the Supreme Court has not (yet) formally endorsed this collapse,[26] the case law often treats these standards loosely, as a general requirement of reasoned fact-finding and decisionmaking, and I will do the same.

In this legal setting, the views of an expert panel will be part of the information on which the agency bases its factual findings. Under the *Universal Camera Corp. v. NLRB*[27] decision(s), the reviewing court is not disabled from looking behind the agency's findings to the

---

20. 5 U.S.C. § 706(2)(A).

21. *Id.* § 706(2)(E).

22. Consol. Edison Co. v. NLRB, 305 U.S. 197, 229 (1938).

23. KEITH WERHAN, PRINCIPLES OF ADMINISTRATIVE LAW 320 (2008).

24. Ass'n of Data Processing Serv. Orgs. v. Bd. of Governors, 745 F.2d 677, 683 (D.C. Cir. 1984).

25. *See* RICHARD J. PIERCE, JR., SIDNEY A. SHAPIRO & PAUL R. VERKUIL, ADMINISTRATIVE LAW AND PROCESS 367–69 (3d ed. 1999); WERHAN, *supra* note 23, at 325 n.27 (collecting cases).

26. *See* Dickinson v. Zurko, 527 U.S. 150, 158 (1999) (noting the question but not resolving it). *But see* Am. Paper Inst. v. Am. Elec. Power Serv. Corp., 461 U.S. 402, 412 n.7 (1983) (describing the arbitrary and capricious test as more lenient than the substantial evidence test).

27. Universal Camera Corp. v. NLRB, 340 U.S. 474 (1951), *remanded to* 190 F.2d 429 (2d Cir. 1951).

views of the expert factfinder, such as an administrative law judge.[28] Rather the reviewing court can consider the expert findings as part of the record, and overturn the agency's findings if the agency has inadequate reason for rejecting the view of an expert, where the agency and expert disagree.[29] Subsequently, I will argue that, under the best reading of relevant organic statutes, and the best conception of the interaction between those statutes and the APA's standards for judicial review, agencies should be obliged to give a particular type of reason—a second-order epistemic reason—for rejecting the views of an expert panel on matters of fact, causation, or prediction. Unless the agency can give valid reasons to doubt the epistemic quality of the panel's conclusions, those conclusions should outweigh the agency's own judgment.

How do the APA's standards apply under the law? What do they entail in practice, either when an expert panel reaches consensus, or when it has a majority and minority view? At least when an expert panel is equally divided, the prevailing rule is that the agency can adopt either view. In *Oregon Natural Resources Council v. Daley*,[30] the National Marine Fisheries Service appointed a panel of sixteen biologists to review scientific data and determine various factual and predictive questions bearing on whether the west coast coho salmon was a threatened species under the Endangered Species Act. As to a crucial issue, the panel split 8 to 8. The court accepted, as a settled proposition, that "in the event of a scientific disagreement between experts, the [agency] is free to rely on the expert opinion of [its] choice."[31]

The precisely equal split on the panel makes the case somewhat unusual, however. More often, where experts disagree, a panel contains a clear majority view and a dissenting view. Can an agency appeal to the majority of an expert panel, or the majority view among experts in the field, as support for a determination of fact or of causation? I take up this question at length in Parts II and III. The short answer is that the cases are somewhat schizophrenic. When

28. *Id.* at 488.

29. For an example from the D.C. Circuit discussing *American Farm Bureau Federation v. EPA*, No. 06-1410, 2009 WL 437050 (D.C. Cir. Feb. 24, 2009) (per curiam), see *infra* notes 85, 91–93.

30. Or. Natural Res. Council v. Daley, 6 F. Supp. 2d 1139 (D. Or. 1998).

31. *Id.* at 1159; *see also, e.g.*, Fed. Power Comm'n v. Fla. Power & Light Co., 404 U.S. 453 (1972) (setting out this general principle and deferring to Federal Power Commission finding where experts gave opposing testimony).

agencies are quite formal about counting expert noses, as in the black lung cases I will discuss, courts sometimes react in horror, saying that nose counting is no substitute for reasoned decisionmaking. Informal nose counting, however, goes on all the time. The cases are legion in which agencies or, for that matter, reviewing courts appeal vaguely to expert consensus or to what most experts or a majority of experts hold. As I will argue, there is a good epistemic reason for that practice.

Must an agency follow expert consensus or even the view of a majority of experts? As I mentioned, some organic statutes, notably the Clean Air Act (CAA), establish an expert panel and require the agency to submit its proposed rules to the panel. Although the agency is not obligated by statute to do what the panel says, the agency must give an adequate reason if it rejects the findings or recommendations of a panel majority.[32] Sometimes, courts unthinkingly construe a requirement of reason-giving as merely requiring a first-order reason—a reasoned statement of the agency's substantive views about the factual, causal, or predictive questions at issue.[33]

I will argue that this is mistaken, and that courts should require the agency to give no less than a valid second-order reason: a valid argument that the expert panel's factual findings are epistemically suspect, relative to that of the agency. Absent such a second-order reason, the views of the expert panel as to factual matters should trump the agency's own first-order judgment, which represents nothing more than another expert vote and, as such, is defeated by the expert consensus or (super)majority view to the contrary, when that exists. Nothing in my analysis, however, implies that agencies must defer to a panel's view about how alternative policies should be evaluated, given the facts. As to that question, the agency's judgment controls, insofar as the law permits. The consequence is that the agency will have an incentive to be clear when its decisions are based on evaluative differences with the panel, thus making the agency's normative commitments transparent to outside monitors—namely, Congress, courts, and the public.

Interestingly, although the voting rules for such panels are usually left unclear, a simple majority of experts, rather than a supermajority, is usually assumed to be controlling, not in the sense that it can bind the agency without more, but in the sense that a

32. 42 U.S.C. § 7607(d)(3) (2006).

33. *See, e.g.*, Stalcup v. Peabody Coal Co., 477 F.3d 482, 484 (7th Cir. 2007).

majority is entitled to state the view of the panel. As we will see, there is a sound epistemic reason for this assumption. Under a broad range of conditions, the views of a simple majority of experts are more likely to be correct than the views of any other subgroup; requiring a supermajority of experts on the panel to agree in order to force the agency to respond with reasons would, effectively, privilege the views of a lesser subgroup with bad epistemic consequences.

II. WHEN AGENCIES MAY COUNT NOSES

If the weight of expert opinion on the panel is *X*, is that an adequately reasoned basis for the agency to find that *X* is the case? For purposes of administrative law, the question is whether nose counting amounts to nonarbitrary fact-finding in cases off the record, or whether nose counting provides substantial evidence in cases on the record. As these standards largely overlap, I will refer, for brevity, to the agency's obligation to engage in reasoned fact-finding.

A. An Example: Black Lung Benefits

Consider the dilemma that faces administrative law judges deciding cases under the Black Lung Benefits Act.[34] Roughly, the statute gives coal miners who suffer from pneumoconiosis, or black lung disease, a claim for benefits against the employer if the miners or their survivors can show, among other things, that the miner contracted pneumoconiosis and that the condition caused death or disability.[35] Administrative law judges, often faced with conflicting diagnoses from doctors who specialize in black lung problems, constantly attempt to resolve the cases by counting noses. And the courts repeatedly rebuke them for doing so.

In a rather typical case, *Stalcup v. Peabody Coal Co.*,[36] five doctors offered opinions on whether the claimant had pneumoconiosis. The administrative law judge held that, because the five doctors were equally qualified, numbers should prevail; because three of the five had found that pneumoconiosis was not present, the finding went against the claimant.[37] The Seventh Circuit said that this was unreasoned decisionmaking. Because a "scientific dispute must

34. 30 U.S.C. §§ 901–45 (2006).
35. *Id.* § 901(a).
36. Stalcup v. Peabody Coal Co., 477 F.3d 482 (7th Cir. 2007).
37. *Id.* at 484.

be resolved on scientific grounds,"[38] the court underscored that administrative law judges "must have a medical reason for preferring one physician's conclusion over another's."[39] The judge:

> cannot avoid the scientific controversy by basing his decision on which side had more medical opinions in its favor. This unreasoned approach, which amounts to nothing more than a 'mechanical nose count of witnesses,' would promote a quantity-over-quality approach to expert retention, requiring parties to engage in a race to hire experts to ensure victory.[40]

The concern expressed in the last clause, about the effects of different decisionmaking strategies where the number of experts is up for grabs, is a useful and potentially valid point that I will take up in Part III below. The rest of the analysis is erroneous, although, to be fair, the decision was for the most part following established circuit precedent.[41] It is too demanding to require the administrative law judge to give a first-order medical reason sufficient to arbitrate between the conflicting diagnoses of two groups of experts in the field. "Avoiding the scientific controversy" is not a moral failing on the part of the lay decisionmaker faced with expert disagreement; it is a perfectly sensible epistemic strategy for lay decisionmakers who lack first-order competence. The administrative law judges' medical conclusions lack any firm epistemic basis, whereas counting the noses of experts has a clear epistemic rationale.

B. Why Count Noses?

The affirmative basis for counting noses is the theory of decisionmaking under risk and uncertainty. Optimal decisionmaking requires optimal information gathering. The agency's problem is to find facts correctly, but it cannot invest unlimited resources in doing so. By counting the noses of experts, under certain conditions, the agency will be able to maximize the overall quality of its decisions, taking into account both the accuracy of its decisions and the costs of decisionmaking.

---

38. *Id.* (quoting Peabody Coal Co. v. McCandless, 255 F.3d 465, 468 (7th Cir. 2001)).

39. *Id.* (quoting *McCandless*, 255 F.3d at 469).

40. *Id.* (quoting Sahara Coal Co. v. Fitts, 39 F.3d 781, 782 (7th Cir. 1994)).

41. *See, e.g.*, Livermore v. Amax Coal Co., 297 F.3d 668, 672 (7th Cir. 2002); *McCandless*, 255 F.3d at 468–69; *Fitts*, 39 F.3d at 782; Old Ben Coal Co. v. Battram, 7 F.3d 1273 (7th Cir. 1993).

What are those conditions? In particular, why should a vote of experts be thought more epistemically reliable than the agency's own views? The answer lies in the rational choice theory of committee decisionmaking. For present purposes, this body of theory can be divided into two branches, involving the *aggregation of information* that panel members possess and the *acquisition of information* by those members.

I will begin with the aggregation question, best approached through the lens of the Condorcet Jury Theorem. Roughly, assume a group of sincere voters—voters trying to get the right answer—whose competence is, on average,[42] at least slightly better than random. With two choices,[43] the voters are slightly more likely to be correct than incorrect. In order to bracket the issue of information acquisition, assume for the moment that the voters' level of competence is exogenous; it is just there.

The Jury Theorem then shows that a majority vote of this group is increasingly likely to be correct as the size of the group increases, as the average competence of its members increases, or as its cognitive diversity increases, when diversity means that the biases of the group's members are negatively correlated.[44] When the voters are themselves experts, the second two conditions—competence and diversity—tend to work at cross-purposes. Experts tend to have high individual competence, but may also have highly correlated biases, perhaps because of common professional training or because they copy each other's opinions. I return to these issues shortly.

For now, the key point is just that majority voting most effectively aggregates the information dispersed among the panel of experts. Nose counting of the assembled experts is a means by which the agency can in effect aggregate expert views, even if the agency itself lacks first-order competence. The conceptual mistake in critiques of nose counting is the idea that, if the agency lacks first-order reasons for its findings, it must be acting in an unreasoned way.

---

42. *See* Bernard Grofman, Guillermo Owen & Scott L. Feld, *Thirteen Theorems in Search of the Truth*, 15 THEORY & DECISION 261, 273–74 (1983).

43. The Theorem can be extended to more than two choices, but as nothing in my discussion depends on this wrinkle, I will assume the two-option case. Note that at least some multiple-option cases can be reduced to two-option cases through successive pairwise comparison.

44. Lu Hong & Scott E. Page, Some Microfoundations of Collective Wisdom 7 (May 12, 2008) (unpublished manuscript), *available at* http://telechargeu.cines.fr/3517/load/documents//cerimes/UPL30290_Page.pdf.

But if the agency's nose counting is itself a rational second-order epistemic strategy, the critique collapses.

The Jury Theorem can be extended to qualified majority, or supermajority rules, but only with restrictions.[45] It has been shown that qualified majority rules maximize the probability of making a correct decision, but only if the status quo is stipulated to prevail in the event that no alternative garners the requisite supermajority.[46] This is a suspect condition in the administrative state, where the regulatory status quo—which may just be the default common law baseline—has no necessary priority, either in theory[47] or under the terms of the APA.[48] If the status quo preference is abandoned, then a weaker result holds: "for sufficiently large [decisionmaking groups] . . . if the average competence of the voters is greater than the fraction of the votes needed for passage . . . a group decision is more likely to be correct than the decision of a single randomly chosen individual."[49] This condition is stringent; if the decisionmaking group uses a two-thirds majority rule, for example, then average competence must be at least .67. Absent these conditions, the background logic of the Jury Theorem is that majority voting is epistemically preferred. Any lesser subgroup of decisionmakers is less likely to be correct, given the Theorem's other conditions; and majority voting alone gives no privilege to the status quo, in line with the administrative state's general assumption that failure to regulate when regulation is justified is as dangerous as unjustified regulation.

The problem of acquiring information merits a brief discussion. The standard treatment of the Jury Theorem assumes that information is exogenous. If this assumption is relaxed, it is apparent that there is a tradeoff between the number of experts and the epistemic quality of their views. If information is exogenous, then the more experts, the more likely it is that the group decision will be

45. This paragraph is adapted from Adrian Vermeule, *Many-Minds Arguments in Legal Theory*, 1 J. LEGAL ANALYSIS 1 (2009).

46. Ruth C. Ben-Yashar & Shmuel I. Nitzan, *The Optimal Decision Rule for Fixed-Size Committees in Dichotomous Choice Situations: The General Result*, 38 INT'L ECON. REV. 175, 179–83 (1997).

47. Cass R. Sunstein, *Reviewing Agency Inaction After* Heckler v. Chaney, 52 U. CHI. L. REV. 653, 656–57 (1985).

48. APA provisions treat agency action and inaction equivalently. 5 U.S.C. §§ 551(13), 706(1) (2006). In practice, however, reviewing courts are more reluctant to force agency action than to block it. *See, e.g.*, Heckler v. Chaney, 470 U.S. 821, 831 (1985).

49. Mark Fey, *A Note on the Condorcet Jury Theorem with Supermajority Voting Rules*, 20 SOC. CHOICE & WELFARE 27, 31 (2003).

accurate, so long as each is at least slightly better than random and holding all else constant. However, when experts must decide how much information to acquire—how much epistemic effort to invest in the panel's activities—increasing numbers make each expert less likely to be the decisive vote, which reduces the effort each will expend. Increasing panel size need not increase the aggregate information panelists hold, because less information will be acquired.[50] In short, with endogenous information, experts have an incentive to engage in epistemic free-riding.[51]

This problem has at least three consequences. First, it creates an optimization problem: institutional designers setting up committees must trade off the quantity of panelists against the quality of their contributions and votes.[52] In most of the real-world cases I discuss, however, Congress has set or capped the number of panel members, or the agency has done so before the agency action giving rise to the litigation, or the number is extrinsically determined in some other way. I will thus bracket this set of issues. Second, the danger of epistemic free-riding within the panel may give agencies good second-order reasons to reject the panel's recommendations, under conditions where the problem is especially likely to be serious. I return to these problems below.

Finally, the problem of endogenous information provides further reason to think that majority rule is the best voting rule for expert panels and the best nose counting rule for agencies attempting to take advantage of expert opinion, at least where experts face highly complex regulatory problems. It has been shown that, as the quality of information decreases, so that each voter will get a very imperfect idea of the truth even after investing effort, the optimal voting rule will fall from unanimity down toward simple majority.[53] The basic intuition is that the larger the supermajority needed to make a decision, such as making a finding for the agency to use, the less each panelist will invest in acquiring information, because the imperfection

50. For this reason, polling the members of the National Academy of Sciences about a relevant problem might well yield less information than voting within a small panel of experts focused on the same problem. Thanks to Stuart Benjamin for raising this issue.

51. Christian List & Philip Pettit, *An Epistemic Free-Riding Problem?*, *in* KARL POPPER: CRITICAL APPRAISALS 128, 138–40 (Philip Catton & Graham Macdonald eds., 2004).

52. Drora Karotkin & Jacob Paroush, *Optimum Committee Size: Quality-Versus-Quantity Dilemma*, 20 SOC. CHOICE & WELFARE 429, 433 (2003).

53. Nicola Persico, *Committee Design with Endogenous Information*, 71 REV. ECON. STUD. 165, 167 (2004).

of the information acquired makes it less likely that any individual panelist will be decisive. (Imagine sitting on a twelve-member panel that requires a unanimous vote to act. A vote only matters if all eleven others vote identically. If the facts are highly unclear, what are the chances that everyone will vote the same way?) If the panel faces an extremely blurry informational environment, as it does in most of the hard regulatory problems for which Congress or the agency has thought an expert panel necessary, then a voting rule approaching simple majority is likely to be best.

A similar point holds if the concern is not that experts fail to invest in acquiring information, but rather that experts will manipulate or distort the information they already possess in order to produce preferred outcomes. In one illuminating model of a deliberative expert committee,[54] majority rule induces the members to reveal their private information with less distortion than under unanimity.[55] The reason is that if a unanimous vote is necessary to depart from the status quo, then members biased in favor of change have strong incentives to overclaim or otherwise manipulate their information. By contrast, majority rule minimizes the net incentives for distortion by panel members with different biases for and against change.

C. Alternatives

Nose counting is hardly the only second-order epistemic strategy that the lay decisionmaker can use when confronted with disagreement among experts. Nose counting is just one such strategy among many. The decisionmaker can examine the relative qualifications of the experts, giving greater weight to the more highly qualified; the *Stalcup* court's opaque preference for a qualitative approach suggests this,[56] and other cases are more explicit that qualification weighing counts as adequately reasoned decisionmaking in black lung cases where expert diagnoses disagree.[57] Another

54. David Austen-Smith & Timothy J. Feddersen, *Deliberation and Voting Rules*, *in* SOCIAL CHOICE AND STRATEGIC DECISIONS 269 *passim* (David Austen-Smith & John Duggan eds., 2005); *see also* David Austen-Smith & Timothy J. Feddersen, *Deliberation, Preference Uncertainty, and Voting Rules*, 100 AM. POL. SCI. REV. 209, 210 (2006).

55. Austen-Smith & Fedderson, *Deliberation and Voting Rules*, *supra* note 54, at 273.

56. Stalcup v. Peabody Coal Co., 477 F.3d 482, 484 (7th Cir. 2007).

57. *See, e.g.*, Old Ben Coal Co. v. Battram, 7 F.3d 1273, 1278 (7th Cir. 1993) ("The ALJ must attempt to evaluate opinions by considering . . . the qualifications of the experts . . . and any other relevant evidence. Because the ALJ . . . considered these factors . . . we may accept his

approach is to examine the track records of different experts to see whether their diagnoses were later proven true or false. Finally, the decisionmaker can discount expert views by the known or apparent biases of experts, who may have self-interested reasons for opining one way or another. Where a company doctor diagnoses no disease, and an independent doctor appointed by the court finds to the contrary, the result is not a tie.

But the fact that nose counting is only one second-order epistemic strategy for choosing among the conflicting views of experts does not show that there is anything wrong with nose counting. Depending upon the situation, it may be a perfectly sensible approach, and indeed perhaps the only available approach. On the facts of *Stalcup*, for example, the other strategies were unavailable. The experts were equally qualified, and there is nothing in the reported facts suggesting that some had better track records than others. In regulatory domains of higher uncertainty and complexity than the black lung cases, experts are especially likely to lack clear track records because it will usually be unclear whether experts' past causal judgments and predictions were accurate.

In some cases, then, there is no real alternative to nose counting. This is not an unusual situation; it is a chronic condition in lay decisionmaking on subjects as to which expertise is relevant, but experts disagree. The lay decisionmaker knows that the majority of experts might be wrong, but placing one's epistemic bets with the majority is still better than placing them with the minority. Implicit nose counting thus goes on all the time. When a patient gets a second opinion that disagrees with the first, a common recourse is to seek a third opinion to break the tie. Captains of the age of sail would bring three compasses to sea; in case the readings given by the first two conflicted, the third would be consulted.[58] And judicial opinions themselves constantly refer, in an untheorized way, to the views of "a majority of experts" or "a majority of studies" to buttress their factual claims.[59]

[conclusion]."); Adkins v. Dir., Office of Workers' Comp. Programs, 958 F.2d 49, 52 (4th Cir. 1992) ("[T]he ALJ must give some reasoned explanation why [the expert's] superior qualifications do not carry the day.").

58. I thank Jon Elster for making this point to me.

59. *See, e.g.*, J.E.B. v. Alabama *ex rel.* T.B., 511 U.S. 127, 138 n.9 (1994) ("majority of studies"); John M. v. Stone, 72 F. Supp. 2d 316, 322 (S.D.N.Y. 1999) ("majority of experts"); United States v. Galbreth, 908 F. Supp. 877, 892 (D.N.M. 1995) ("majority of experts").

Indeed, implicit nose counting also underpins the common but untheorized agency practice in which agencies appeal to expert consensus as a sufficient reason for taking facts as established. Such an appeal effectively says that the unanimous view of experts is sufficient reason to take a fact or causal theory as established; this is just as much a case of nose counting as an appeal to a majority view of experts. But if first-order reasons were the only permissible reasons, the difference between unanimity and majority should make no difference. The logic of cases like *Stalcup* is that administrative decisionmakers should not be able to appeal to expert consensus either—they should have to disgorge their first-order judgments, whether or not the experts have achieved consensus—but this is preposterous.

D. Nose Counting in Regulatory Rulemaking

The black lung cases involve administration of a benefits scheme. In marked contrast to those cases, in which nose counting is said to be impermissible, nose counting by both agencies and reviewing courts is ubiquitous in regulatory rulemaking. The practice, however, is almost wholly untheorized.

In *American Trucking Ass'ns v. EPA*,[60] the D.C. Circuit reviewed EPA's 1997 revision of the NAAQS for ozone and particulate matter. In an initial round of litigation, the D.C. Circuit had invalidated the rulemaking on constitutional nondelegation grounds, only to be reversed by the Supreme Court.[61] On remand, the D.C. Circuit upheld the rules under standard arbitrariness review. Notably, EPA had justified its choices not only by first-order reasons, but also by nose counting. One key feature of the rules, EPA pointed out, was within a range thought desirable by "most CASAC . . . members."[62] As to another key feature, "EPA . . . emphasized that CASAC unanimously agreed with the proposed change."[63] The court followed suit, upholding the EPA's decisions in part because the court thought that the agency's nose counting was a perfectly sound sort of reason for the agency to give. As to one issue, the court specifically cited EPA's appeal to CASAC consensus, calling it "record evidence" that helped

60. Am. Trucking Ass'ns v. EPA, 283 F.3d 355 (D.C. Cir. 2002).

61. Am. Trucking Ass'ns v. EPA, 175 F.3d 1027 (D.C. Cir. 1999), *rev'd sub nom.* Whitman v. Am. Trucking Ass'ns, 531 U.S. 457 (2001).

62. *Am. Trucking Ass'ns*, 283 F.3d at 367.

63. *Id.* at 376.

the agency meet its burden of reasoned fact-finding and decisionmaking.[64] As to another issue, on which CASAC did not reach consensus, the court noted that EPA had followed the majority view.[65]

The opinion on remand in *American Trucking* is eminently sensible, and it exemplifies an epistemic strategy that is routine for agencies and courts. Current administrative law is inconsistent about whether nose counting is an adequate basis for fact-finding for purposes of APA review, with the black lung cases saying that it is not and cases like *American Trucking* suggesting that it is. I believe the second view is generally correct.

III. WHEN NOSE COUNTING IS INADEQUATE

Despite these points, and compatible with them, it is also true that nose counting is inadequate under a narrow set of identifiable conditions, and should be ruled out as a valid basis for agency fact-finding and decisionmaking. I will identify two such conditions: when nose counting would create incentives for costly strategic behavior by litigants or agencies, and when there is a demonstrable risk of biased decisionmaking or harmful groupthink on expert panels. Let us discuss these in turn.

A. Endogenous Panels and Strategic Behavior

The *Stalcup* court offered the valid concern that nose counting "would promote a quantity-over-quality approach to expert retention, requiring parties to engage in a race to hire experts to insure victory."[66] Where the very size of the panel is itself a variable that the parties or the agency can choose, and can choose during litigation or in the shadow of litigation, then the number of noses to be counted is endogenous. In the worst version, parties or agencies will add experts after seeing what earlier experts have said, in order to generate a desired result.

Such concerns, however, need not entail that nose counting is altogether impermissible. In many, perhaps most, cases of agency nose counting, the panel's number and composition is set or constrained by statute, agency rule, or policy well before agencies

64. *Id.* at 378.

65. *Id.* at 379.

66. Stalcup v. Peabody Coal Co., 477 F.3d 482, 484 (7th Cir. 2007).

have found relevant facts in support of a proposed rule or order, and before litigation ensues. Under standard principles of administrative law, an agency directive creating a panel is binding upon the agency unless and until changed through proper procedures.[67] And, as a backup safeguard, courts can apply greater scrutiny to panels whose number or composition the agency has determined after litigation has commenced.[68] In such cases, the scope for strategic behavior is small or nonexistent; the panel can safely be treated by reviewing courts as entirely exogenous.

What makes the problem a real one in the black lung setting, and in other settings in which agencies award statutory benefits or impose statutory penalties conditioned on adjudicative facts, is that the parties themselves hire experts after seeing what the other side's experts have said, or what their own experts have said, and after seeing how many experts the other side has hired. In such settings, rational parties will invest in hiring additional experts up to the point at which the marginal increase in the expected payoff equals the marginal cost of hiring the next expert. There is also a serious risk that later experts will simply free-ride on the conclusion of earlier experts, as I will discuss shortly.

Here too, however, making nose counting impermissible is unnecessary. The simpler approach is to require parties to use a standard operating procedure, familiar from arbitration, in which the number of experts is determined in advance, behind a veil of ignorance about what the experts will say. Whether agency adjudicators or reviewing courts have the authority to do this will of course depend upon the details of the statutory scheme, but there is no objection in principle. Here the core problem is not with nose counting as such, but with strategic behavior. The problem can be attacked on its own terms, without an overbroad prohibition on an epistemically sensible strategy of second-order reasoning.

Moreover, the *Stalcup* court's concern about strategic behavior, however valid in itself, just identifies a collateral systemic cost that can arise from nose counting. It is a separate question whether that cost outweighs the epistemic benefits. A system or practice with both

67. *See* Ariz. Grocery Co. v. Atchison, Topeka & Santa Fe Ry. Co., 284 U.S. 370, 389–90 (1932).

68. *Cf.* Motor Vehicle Mfrs. Ass'n v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 50 (1983); NLRB v. Metro. Life Ins. Co., 380 U.S. 438, 444 (1965); SEC v. Chenery Corp., 332 U.S. 194, 196 (1947).

nose counting and some degree of strategic behavior may be superior overall to a system or practice with neither feature. If the agency has a reason for thinking that the benefits outweigh the costs, that reason is itself subject to the APA's usual standards of review, and courts should uphold it so long as it is nonarbitrary.

Finally, I speculate, without hard evidence, that nose counting is impossible to eliminate altogether. Agencies or agency adjudicators who are basically at sea—who rationally lack any confidence in their own first-order judgments, for example, the sort of medical judgments that the *Stalcup* court oddly required agency adjudicators to give—may implicitly count noses for lack of a better alternative, as indeed judges sometimes do as well. In such cases, making nose counting impermissible will simply drive agencies to conceal the real bases for their fact-finding. Nose counting is a ubiquitous epistemic strategy, both in ordinary life and in official decisionmaking, and administrative law has no real choice but to accommodate it.

B. Judgment Falsification and Groupthink

Part of the judicial concern about nose counting may rest on an intuitive concern about false consensus, herding, or groupthink by experts. In a Jury Theorem framework, the concern is that a group of experts in a given field will have highly correlated biases, because of common professional training, because they copy each other's opinions, or even because the expert panel fakes an appearance of consensus for public consumption; in the last case, experts will not be voting sincerely, which undermines the operation of the Theorem.

Common professional training is a built-in hazard of expert panels drawn from a scientific field or subfield. Copying may occur because of an "information cascade," in which individual experts rationally use the views of other experts as the basis for forming their own views, thus reducing the number of independent opinions expressed by the group overall. Here there is a kind of epistemic free-riding[69] or "cognitive loafing,"[70] as some within the group benefit from the information provided by other's views without contributing information themselves. Copying may also occur because of a

69. *See* List & Pettit, *supra* note 51, at 138–40.

70. Mark Seidenfeld, *Cognitive Loafing, Social Conformity, and Judicial Review of Agency Rulemaking*, 87 CORNELL L. REV. 486, 486 (2002).

"reputational cascade,"[71] in which experts follow the views of senior scientists or powerful figures in the field, for fear of being stamped as incompetent or odd.

As to the falsification of expert judgments, case studies have shown that expert panels sometimes gin up a consensus that does not actually exist.[72] The panel may do this in order to maximize its members' joint influence on agencies and other decisionmakers, or out of paternalistic concern that the agency or public will become confused if the panel ventilates its disputes, or because the panel's members have a professional interest in preserving a public reputation for expertise.[73] When this occurs, some panel members are falsifying their judgments, and the panel as a whole conceals information—about the presence and magnitude of expert disagreement—that is useful for decisionmakers.

C. The Conditions for Expert Groupthink

Whether such concerns are serious depends upon the composition, structure, and decisionmaking process of the expert panel. What factors make groupthink or judgment falsification more or less likely? I will use groupthink, itself an ill-defined notion, as shorthand for the various forms of epistemic free-riding, informational and reputational cascades, and falsification of judgments I have mentioned. Although the problems are somewhat different, the institutional determinants of the various types of groupthink overlap a great deal.

First, groupthink is less likely to occur as the diversity of panel membership increases.[74] Many panels are chosen uniformly from specialists in relevant scientific subfields, like the panel of scientists in *Oregon Natural Resources Council v. Daley*.[75] In other cases, statutes require that expert panels contain professionals from different fields, or even nonprofessionals,[76] and these can be understood as means for diversifying the panel's training, assumptions, and intellectual

71. Timur Kuran & Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 STAN. L. REV. 683, 685–89, 727–28 (1999).

72. John Beatty, *Masking Disagreement Among Experts*, 3 EPISTEME 52, 55 (2006).

73. *Id.* at 53–54. *See generally* Bauke Visser & Otto H. Swank, *On Committees of Experts*, 122 Q.J. ECON. 337 (2007) (showing conditions under which a panel of experts concerned for their individual reputations will generate false consensus).

74. CASS R. SUNSTEIN, WHY SOCIETIES NEED DISSENT 141–44 (2003).

75. Or. Natural Res. Council v. Daley, 6 F. Supp. 2d 1139, 1146 (D. Or. 1998).

76. *See infra* notes 116–120 and accompanying text.

outlook. In Jury Theorem terms, this approach trades off reduced average competence, because diversification requires that some specialists in the subfield must be bumped off a panel of fixed size, for reduced correlation of biases across the group. Depending upon the precise composition, the benefits of reduced correlation can more than compensate for the loss, in which case some degree of diversification will be epistemically optimal.[77] I return to these issues shortly.

Another major cause of groupthink is sequential, rather than simultaneous, expression of views among the panel experts. If experts express their judgments in ignorance of other experts' judgments, herding and cascades are ruled out, although false consensus arising from experts' concern for reputation is still possible.[78] Ideally, experts should vote simultaneously rather than sequentially, in order to prevent informational and reputational cascades. In real-world conditions, however, simultaneity is difficult to achieve; deliberation prior to voting will give experts a sense of where other experts stand. These tensions are on display in a set of guidelines issued by the Food and Drug Administration for its many expert advisory committees.[79] On the one hand, the guidelines expressly recommend simultaneous voting, citing the academic literature on the risks of information cascades.[80] On the other hand, the guidelines recommend extensive deliberation before voting, and recommend against the use of secret ballots,[81] which can help block reputational cascades by preventing panel members from knowing how others voted.

The structure of experts' compensation is also important. If experts receive no compensation or merely nominal compensation, as is the case with many scientific panels, then the incentives for

77. Krishna K. Ladha, *The Condorcet Jury Theorem, Free Speech, and Correlated Votes*, 36 AM. J. POL. SCI. 617, 629 (1992).

78. For a model in which expert panels, voting simultaneously, nonetheless generate false consensus, see Visser & Swank, *supra* note 73. The basic mechanism in the model is that experts believe that the audience believes that competent experts will all have the same view of the facts, in which case disagreement among the panel implies that some of its members are less competent. On further assumptions, this causes the minority to go along with the majority even if the minority disagrees. This mechanism is unaffected by the simultaneity of voting.

79. *See* U.S. DEP'T OF HEALTH & HUMAN SERVS., FOOD & DRUG ADMIN., GUIDANCE FOR FDA ADVISORY COMMITTEE MEMBERS AND FDA STAFF: VOTING PROCEDURES FOR ADVISORY COMMITTEE MEETINGS (2008), *available at* http://www.fda.gov/oc/advisory/GuidancePolicyRegs/ACVotingFINALGuidance080408.pdf.

80. *Id.* at 5 n.1.

81. *Id.* at 4.

epistemic free-riding or cognitive loafing are at a maximum, increasing the risks of herding within the panel. Even if experts are rewarded for their own performance, information cascades are still a risk, because in an information cascade, it is individually rational to follow the views of a sufficient number of others, even if one's private information is to the contrary. Both theory[82] and experiments[83] suggest that the best way to prevent information cascades from forming is to reward individuals on the basis of the accuracy of a majority vote within the group, rather than for individual accuracy. Under this reward structure, the individual's incentive is to reveal private information to the group, maximizing the chance that (a majority of) the group as a whole will make the correct decision.

All this said, however, the groupthink concern does not show that nose counting is necessarily an impermissible epistemic strategy for agencies. What it shows is that agencies and reviewing courts will sometimes have valid second-order, epistemic reasons for discounting the views of an expert consensus or an expert majority. I turn to such cases in Part IV. In other cases, however, such concerns are not implicated; or, if they are implicated, the agency may still rationally decide that nose counting is a better epistemic strategy than the available alternatives. Where either of these conditions is met, there is nothing wrong with nose counting of experts by agencies.

D. Conclusion

I conclude that a factual finding based upon nose counting of experts should count as an adequately reasoned decision, absent special reason for second-order concern about strategic behavior, judgment falsification, or expert groupthink. When those conditions do not hold, nose counting is a valid basis for decision within the terms of the APA.[84] The court's basic error in cases like *Stalcup* is to think that first-order reasons—the administrative law judge's medical conclusions, issued in the face of expert disagreement—are the only

82. Vladislav Kargin, *Prevention of Herding by Experts*, 78 ECON. LETTERS 401, 402 (2003).

83. Angela A. Hung & Charles R. Plott, *Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions*, 91 AM. ECON. REV. 1508, 1509 (2001).

84. *Cf.* Fed. Power Comm'n v. Fla. Power & Light Co., 404 U.S. 453, 464–65 (1972) ("[W]ell-reasoned expert testimony—based on what is known and uncontradicted by empirical evidence—may in and of itself be 'substantial evidence' when first-hand evidence on the question . . . is unavailable.").

sort of reasons that administrative law can recognize. Nothing in the Act, in the conceptual framework of administrative law, or in the theory of rational decisionmaking requires this exclusion of second-order reasons. Nose counting of experts is a ubiquitous second-order strategy for lay decisionmakers; although it is not the only such strategy, there is often no alternative. To force lay decisionmakers, in such cases, to arbitrate an expert disagreement by coughing up a first-order reason for which they lack any epistemic foundation is itself a guarantee of unreasoned decisionmaking.

IV. WHEN AGENCIES MUST COUNT NOSES

So far we have discussed cases in which agencies use nose counting as a reason for their decisions, and courts review whether the reason is adequate. By contrast, there are also important cases in which agencies wish to depart from the majority or supermajority view of an expert panel, as to matters of fact, causation, or prediction. Should they be allowed to do so? Under what conditions, and based on what reasons?

A. An Example: Fine Particulate Matter

A well-known example of this problem, during the Bush administration, was EPA's 2006 decision to reject the CASAC recommendations about the revision of the NAAQS for fine particulate matter.[85] CASAC, following a 20 to 2 vote of its subpanel on particulate matter, recommended an annual standard between twelve and fourteen micrograms per cubic meter, but EPA rejected the recommendation and maintained the extant annual standard of fifteen micrograms. This was the first time that EPA had ever directly rejected a CASAC recommendation in the NAAQS revision process.

A section of the CAA obligates an administrator to explain why proposed rules "differ[] in any important respect from any of . . . [CASAC's] recommendations."[86] The administrator explained the EPA's 2006 rejection on several grounds: (1) the agency's choice

85. The 2006 NAAQS regulations that resulted were successfully challenged, so far as relevant here, in the D.C. Circuit by a coalition of state attorneys general and health and environmental groups. *See* Am. Farm Bureau Fed'n v. EPA, No. 06-1410, 2009 WL 437050, at *6 (D.C. Cir. Feb. 24, 2009) (per curiam) (overturning the EPA's decision as to the annual standard for fine particulate matter on the ground that EPA failed to adequately explain its decision).

86. 42 U.S.C. § 7607(d)(3) (2006).

was based on its own view of "the most directly relevant body of scientific studies,"[87] (2) CASAC saw less scientific uncertainty than the Administrator thought was actually present,[88] and (3) the CASAC recommendation was not unanimous, because two of twenty-two members of the particulate matter panel dissented.[89] Compressing these grounds somewhat, we may understand the administrator as giving two types of reasons for departing from CASAC's findings and recommendation: (1) first-order scientific reasons based on the agency's own expertise—in this case its expert assessment of the best available science, and of the level of scientific uncertainty—and (2) the second-order reason, offered to diminish the epistemic force of the panel's conclusions—that the panel was not unanimous, implying that reasonable experts could disagree.

I will suggest that neither type of reason was sufficient to justify rejecting the panel's factual findings. The former is just the administrator's first-order view, and has no special epistemic status;[90] the latter is, in these circumstances, an invalid second-order argument. Lacking any valid reason for departing from the panel's findings, the administrator violated the CAA. In order to prevail, the administrator would have had to either give a valid second-order reason for rejecting the panel's factual findings, or else show that the agency's regulatory priorities or evaluation of alternative policies were an adequate basis for disagreement. In the latter case, the administrator would have had to make transparent EPA's normative differences with the panel, making it easier for Congress, the courts, and the public to monitor the agency's commitments and behavior.

As it turned out, the D.C. Circuit recently overturned the administrator's decision as to the annual standard for fine particulate matter, finding that the administrator "failed adequately to explain its reason for not accepting the CASAC's recommendations."[91] The court's analytic framework was somewhat different than the one I

87. National Ambient Air Quality Standards for Particulate Matter, 71 Fed. Reg. 61,144, 61,174 (Oct. 17, 2006) (codified at 40 C.F.R. § 50.6 (2008)).

88. *Id.*

89. *Id.* at 61, 174 n.44.

90. Again, the administrator's superior democratic credentials might entitle him to evaluate alternative policies within the bounds Congress has set out, but they are no basis for according his judgments of fact, causation, or prediction superior status to those of (other) experts. My discussion is confined to the latter point.

91. Am. Farm Bureau Fed'n v. EPA, No. 06-1410, 2009 WL 437050, at *7 (D.C. Cir. Feb. 24, 2009) (per curiam).

suggest because the court said that the administrator's technical expertise deserves deference,[92] but in practice the court seemed to weigh the technical expertise of CASAC more heavily,[93] consistent with the analysis here.

B. Expert Disagreement and First-Order Reasons

Contrary to current law, I suggest that the administrator's first-order view of the scientific merits should have no special weight as against a panel's factual findings. Administrators are experts, at least in the sense that their views are typically informed by the expertise of staff scientists (although as we will see, this is not always true, and was not true in the particulate matter rulemaking). But the members of the scientific panel are expert as well. The administrator's view should thus be understood to count as just another vote among others. If the expert panel is tied, it follows that the agency's vote is decisive and the agency can choose either view, as the court held in *Oregon Natural Resources Council v. Daley*. If, however, a decisive majority of experts favors a certain view, then the administrator's contrary view is simply another vote for the dissenting side, and as such is defeated.[94]

This approach is entirely consistent with believing—if anyone does believe this—that the administrator is the single most competent expert in the picture. A main point of the Jury Theorem is that a group of somewhat less expert voters, so long as their competence is better than random, can be markedly superior to a single expert of higher competence. Indeed, under identifiable conditions, the accuracy of the group's median or mean member will necessarily exceed that of its single most competent member.[95]

The agency can, of course, hear what the panel has to say, review its methods and conclusions, and then form its own judgments. This

---

92. *See id.* at *14. As I explain below, my suggestion is not that the administrator lacks expertise, but that the administrator should be understood to have only one vote in the parliament of the experts.

93. *See id.* slip op. at *7–*10 (noting disagreements between CASAC and EPA, and finding that EPA had not adequately explained why its judgments should trump those of CASAC).

94. This approach creates a discontinuity: where the expert panel is tied the administrator necessarily prevails, but not otherwise, so the switch of a small number of votes on the panel can in theory be dispositive for the outcome. This is, however, a standard property of majority rule—under May's Theorem, it is one of the conditions that majority rule can alone satisfy jointly—and it is a routine property of majority voting in legislatures.

95. SCOTT E. PAGE, THE DIFFERENCE: HOW THE POWER OF DIVERSITY CREATES BETTER GROUPS, FIRMS, SCHOOLS, AND SOCIETIES 158 (2007).

does not show that the agency's conclusions can incorporate all of the panel's expertise, and add to it the agency's own. To see this, imagine that the administrator is placed on the panel; some statutory panels, set up to advise a given agency, do include the heads of other agencies or government officials. In such circumstances, the administrator could likewise consider the information available to other panel members, but the administrator would merely have one vote among others, and a contrary majority view of experts would trump that vote. I suggest that the same logic holds when statutes require the administrator to consider the panel's factual findings and to give a reason for departing from them. In such cases, the agency's own first-order view of the facts should not suffice to trump the panel's aggregate expertise.

C. Ex Ante Incentives

In principle, it is possible that members of expert panels might invest less in acquiring information when, and because, agencies defer to them on factual questions. If the agency must accept the panel's work, the panel might do it less carefully. Yet the opposite effect seems at least equally possible: expert panels who know they will receive no deference have little incentive to get things right, whereas expert panels may invest more time in acquiring information precisely when, and because, they know that the agency is presumptively obliged to accept their findings. Requiring agency deference to panel findings, in other words, eliminates a kind of epistemic moral hazard[96] that can arise when the panel is aware that the agency is likely to ignore its work. In light of this point, the ex ante incentives of the framework suggested here are ambiguous and unclear; at a minimum, there is no ground for concern that the framework will systematically undermine expert panels' incentives to acquire information.

D. Disaggregating Agencies

This analysis assumes that it is best to treat the agency head as an expert, whose view is to count for one but also for no more than one.[97]

---

96. *Cf.* ADRIAN VERMEULE, LAW AND THE LIMITS OF REASON 95–96 (2009) ("Legislators who anticipate constitutional judicial review may rationally invest less in gathering and processing information . . . . Precisely because they know that judges will be trying to catch their mistakes, they may commit more mistakes.").

97. Some agencies have multi-member heads; an example is the five-member Federal Communications Commission. The logic of my suggestion is that each member would count for

I do not, however, defend the very different claim that reviewing courts should ignore internal expert staff who make recommendations to the agency head. Under the *Universal Camera* approach, the findings of internal experts, such as agency technical staff,[98] are themselves part of the record. By parallel, the view of internal expert staff should be taken to be part of a larger, albeit virtual, expert panel composed of both outside expert advisory panels and career experts internal to the agency—the full parliament of experts.

This yields several possible cases. In the easiest case, expert staff within the agency take the same view as outside expert panels, and both are opposed by the agency head, who is a political appointee. During EPA's 2006 revisions of the fine particulate matter standards, the agency's expert staff made findings in accord with CASAC's findings, and the administrator rejected both.[99] Here, adding the votes of the internal experts to those of the expert advisory panel would, at most, increase the size of the expert supermajority opposing the agency's view. That view would be overridden anyway under the approach I have suggested, unless the agency has a valid second-order reason for thinking that the expert panel's conclusions were epistemically informed. Including internal agency staff in the parliament of experts makes no difference in such cases.

There are two harder cases, however. In the first hard case, the internal experts support an agency's view, and both are contradicted by an outside expert panel. Suppose that an expert panel composed of ten members splits, 6 to 4, in favor of a certain finding of fact. Suppose also that the agency head, considering the views of three internal experts, believes the finding is incorrect. Is the count 6 to 5 in favor of the finding, because the agency counts for one? Or is it 8 to 6 against the finding, because one must count not only the agency's own vote, but also the votes of the three internal experts? In the second hard case, an agency view is supported by a majority of the outside

---

one vote in the parliament of experts. Because outside expert panels rarely split by such small margins, this is unlikely to make a difference in many real cases.

98. *See, e.g.*, Natural Res. Def. Council, Inc. v. EPA, 902 F.2d 962, 967–70 (D.C. Cir. 1990) (per curiam) (treating an internal staff paper as relevant to judging the adequacy of the agency's decisionmaking), *vacated in part*, 921 F.2d 326 (D.C. Cir. 1991).

99. Final Opening Brief of State Petitioners and State Amici at 7–8, Am. Farm Bureau Fed'n v. EPA, No. 06-1410, 2009 WL 437050 (D.C. Cir. Mar. 7, 2008) (per curiam), 2008 WL 2609199.

expert panel, but counting the votes of internal expert staff separately would tip the overall count of expert votes against the agency.

The approach I suggest implies that the views of expert staff should be counted in both of these hard cases. Career experts within the agency are at least partly insulated from political interference by civil service protections, thick but informal norms of professional autonomy, and mission orientation.[100] As a result, they not infrequently make findings that politically appointed agency heads override or ignore, as occurred in the particulate matter controversy. Whether they support or oppose the findings of the agency head, the expert views of internal staff provide epistemically valuable information that reviewing courts can and should take into account, without large increases in the costs of judicial decisionmaking.

A cost of this approach is that it might give agencies tactical incentives to hire more internal experts, in the hope of swamping expert advisory panels by sheer weight of numbers. This is a self-limiting problem, however, and thus not a major concern. Hiring internal staff experts is costly business, given the civil service protections and general red tape surrounding federal employment, and the agency will anticipate that it will be saddled with the internal experts even after the particular rulemaking or adjudication at issue is long since past.

E. Facts, Policymaking, and Statutory Authority

The approach I suggest is consistent with the undisputed truth that the administrator, not CASAC, is charged by statute with finding facts and with ultimate legal responsibility for the decision.[101] The problem is to reconcile that authority with the statutory command that the agency must give adequate reasons for rejecting CASAC's findings and recommendations. In the framework I have set out, these provisions are best understood as entrusting the agency with legal responsibility for making the ultimate decision, based in turn on the best available judgments of fact. They entrust the agency with the evaluation of alternative policies, and with deciding what the agency will do given some determination of the state of the world, but they provide no reason to privilege the agency's unilateral determination

100. Sidney A. Shapiro & Richard E. Levy, *Judicial Incentives and Indeterminacy in Substantive Review of Administrative Decisions*, 44 DUKE L.J. 1051, 1064 n.43 (1995) ("[T]he permanent staff of agencies are often resistant to the policy agenda of political appointments.").

101. 42 U.S.C. §§ 7408–09 (2006).

of facts, when the agency's reasons for rejecting expert conclusions are inadequate. The agency is entitled to decide what to do, and can always reject a panel's recommendation on the grounds that the agency evaluates possible policies differently than does the panel, or that the agency wishes to allocate resources differently. The agency, however, is not entitled to make determinations of fact, causation, or prediction that depart from the best epistemic indicator of the truth of those matters—the expert panel's view—absent some second-order reason to think that the epistemic quality of that view was compromised.

In *Universal Camera*, likewise, the Court noted that the agency had been charged with fact-finding and decisionmaking authority, but nonetheless held that on an appropriate record the expert factfinder's conclusions could trump that of the agency. Accordingly it reversed the lower court, which had said that the grant of fact-finding and decisionmaking authority to the agency meant that the agency's findings must prevail over those of a hearing examiner. A main point of *Universal Camera* is that a statutory grant of fact-finding authority to the agency does not immunize the agency's factual conclusions when expert factfinders disagree.

That said, the approach I suggest is squarely inconsistent with the broad principle that, "in the event of a scientific disagreement between experts, the [agency] is free to rely on the expert opinion of [its] choice."[102] This principle is too broad; taken literally, it would allow the agency to rely on the view of even a single dissenting expert, as against the otherwise unanimous view of the whole scientific community. The principle does apply sensibly in cases like *Oregon Natural Resources Council v. Daley*, in which an expert body splits down the middle. When that is so, there is no alternative to letting the agency break the tie. When, however, an expert panel divides into a majority or supermajority and a minority of dissenters, the logic of tiebreaking no longer applies.

F. An Interpretive Default Rule

Congress could of course specify, in the organic statute or elsewhere, that the agency's first-order view of the facts will prevail in the event of a disagreement with a panel consensus or a panel majority; if so, then those first-order views would count as legally

102. Or. Natural Res. Council v. Daley, 6 F. Supp. 2d 1139, 1159 (D. Or. 1998).

sufficient reasons. Congress has not legislated so specifically, however. In the case of the CAA, Congress simply said that the EPA Administrator must supply "an explanation of the reasons for such differences [between CASAC findings or recommendations and EPA's final rules]";[103] so too with the secretary of HHS, who is obliged by statute to give a reason for rejecting a recommendation of the Advisory Committee on Childhood Vaccines (ACCV). It is an open question what sort of reasons such provisions should be understood to require.

Absent more specific direction from Congress, should these provisions and others like them be interpreted to require first-order reasons—substantive views about the matters of fact, causation, and prediction on which CASAC offers its opinions? Or should they be interpreted to require second-order reasons about the relative epistemic quality of the agency's views and CASAC's? As an interpretive default rule, I suggest that, unless Congress clearly does specify otherwise, such provisions should be read to require the agency to give a valid second-order reason for departing from the factual findings of a scientific or technical advisory panel. The best epistemic practice is thus to treat the agency's first-order reasons as just another vote that is outweighed by a contrary body of expert votes on factual matters. The only type of reason that suffices is a reason to think that, epistemically, the agency is better positioned than a (super)majority of the panel to get the facts right. Although such a showing is possible—I canvass some valid second-order reasons that agencies might be able to offer and substantiate in particular cases—it will usually be difficult.

Nothing in this argument, however, suggests that agencies must defer to the panel's evaluation of possible policies, given certain facts. As we have seen, courts allow agencies to set regulatory priorities and allocate resources, even in the face of panel recommendations to the contrary.[104] The logic of that restriction is that agencies may evaluate possible policies in light of their own preferences, rather than the panel's, insofar as the law permits. A statutory duty to give reasons for rejecting a panel's findings and recommendations,[105] then, is best

103. 42 U.S.C. § 7607(d)(3) (2006).

104. *See* Int'l Union, UAW v. Chao, 361 F.3d 249, 256 (3d Cir. 2004) (allowing the agency to reject an advisory committee recommendation on the ground that the agency had rationally set other priorities).

105. 42 U.S.C. § 7607(d)(3). The CAA obligies the Administrator to explain any differences between the adopted policies and CASAC's "findings" or "recommendations." *Id.*

understood as two very different duties with different consequences. As to findings, agencies must adopt the factual bases of a panel's conclusions absent some valid second-order reason for rejecting them. By contrast, the agency need not adopt the panel's overall recommendations for policies if agencies have a standard first-order reason for evaluating competing policies differently or for setting different regulatory priorities. Yet even in such cases, the agency will be forced to openly state its evaluative differences with the panel or its different regulatory priorities, thereby reducing the costs of monitoring to reviewing courts and democratic bodies.

Under ordinary canons of statutory interpretation, I suggest, an approach of this sort makes the best sense overall of the ambiguous relationship between the agency that Congress established and the expert panel that Congress also established. A puzzle arises when the statute itself, as well as background principles of administrative law, make the agency responsible for the decision, yet the statute also gives the panel ability to make recommendations as well as factual findings.[106] In the case of the CAA, for example, there is some tension between Congress's decisions both to make the agency responsible for the ultimate decision and also to specify that EPA must explain its reasons for differing from the panel's findings and recommendations. The most sensible reconciliation, in light of general background principles of administrative law, is that Congress saw the scientific panel as occupying the best epistemic position to make judgments about questions of fact, leaving to EPA the important responsibility of evaluating alternative policies in light of those factual judgments. The same holds for the ACCV and the secretary of HHS.[107]

This is a point about the best interpretation of the agency's organic statute, but it also helps to make sense of relevant APA

106. Thanks to Jonathan Wiener for emphasizing the latter point.

107. In other cases, the organic statute merely creates a scientific or technical advisory committee, but does not in terms require an explanation for differing from its recommendations. *See* Holland-Rantos Co. v. U.S. Dep't of Health, Educ. & Welfare, 587 F.2d 1173, 1175 (D.C. Cir. 1978) (per curiam) ("The recommendations of the NAS-NRC panels are advisory in nature."). Even in such cases, however, the statutory mandate, and the purpose behind it, are most naturally read together to implicitly require that the agency give a valid second-order epistemic reason for departing from the panel's view. *See id.* at 1175–76 (commenting that, although the FDA's "cavalier and unexplained rejection of the opinion of its expert panel strains a 'cornerstone requirement' of the administrative process, that of 'reasoned decision making,'" a remand would be "arid formalism" under the circumstances, because subsequent evidence made it clear that the agency's decision was correct (quoting Columbia Broad. Sys., Inc. v. FCC, 454 F.2d 1018, 1025 (D.C. Cir. 1971))).

standards and their interaction with the organic statutes. The APA's background obligation of reasoned fact-finding and decisionmaking is best understood to require that the agency either defer to the expert panel as to factual matters or else give a reason to think that it, rather than the expert panel, is in the best epistemic position to determine relevant facts. This was the implicit logic of *Universal Camera*, in which the Court said that reviewing judges could look behind agency findings to consider whether the agency had given adequate reason for refusing to credit the contrary findings of a specialized hearing examiner.[108] The Court's discussion, as amplified and clarified by Learned Hand and Jerome Frank in opinions on remand, suggests that examiners are usually best situated to determine witness credibility and demeanor, and that to reject their findings, the agency must give a reason that "results from the [agency's] rational use of the [agency's] specialized knowledge."[109] The agency, in other words, must show that its epistemic competence is greater, on average, than the expert factfinder's epistemic competence, as to the relevant class of questions.

G. Fact-Finding and Second-Order Reasons

On this approach, the key issue, and the key holding of *Universal Camera*, involves *comparative epistemic competence*: whether the agency or expert is best positioned to determine relevant facts, where reviewing courts who lack direct knowledge themselves should place their epistemic bets, and, more generally, how fact-finding authority should be allocated between agencies and their expert advisors. Agencies are not always obliged to find facts in accordance with the expert panel's view. Rather, agencies should be allowed to reject panel findings if they substantiate the right sort of second-order reason.

In general, valid second-order reasons will give reviewing courts confidence that the best epistemic bet is to rely on the agency's first-order views rather than those of the expert panel. Agencies should be able to reject panel findings only if they can give concrete reason to think that the epistemic quality of the expert panel's conclusions are low, relative to the agency's own views. It follows that valid reasons

108. NLRB v. Universal Camera Corp., 190 F.2d 429, 432 (2d Cir. 1951) (Frank, J., concurring).

109. *Id.*

will just be the flipside of the permissible epistemic strategies discussed in Parts II and III.

H. Track Records

In some cases, agencies will have information about the accuracy of an expert panel's past factual estimates, the causal theories it advanced, or its past predictions. When this is so, agencies might be able to show that the panel has often erred. This is likely to be a rare case, however, and courts should be wary of this sort of second-order claim for several reasons.

Mere inaccuracy does not show that a past panel's conclusions were epistemically flawed or that the agency's own first-order views are likely to be systematically superior. Because factual determinations, causal theories, and predictions of any complexity will inevitably have a "stochastic" element—an element of irreducible randomness, arising from the costs of information or built into the nature of things—it is perfectly possible that the panel's conclusions are systematically superior to the agency's, from an ex ante perspective, even if the panel's conclusions have sometimes turned out to be wrong. Furthermore, panels whose members serve for long periods may develop "endogenous expertise"[110] through experience and institutional learning over time. In such cases, the panel's initial findings and predictions may be flawed, but their quality will systematically improve over time. From the standpoint of the reviewing court, the best epistemic bet overall may still be to trust the conclusions of the expert panel over the agency's views.

These points imply that a high rate of turnover among panel members may make it difficult to gauge the track record of the panel. Indeed, as turnover increases, "the panel" becomes an increasingly ill-defined entity. Reviewing courts could plausibly, and without much difficulty, discount the validity of an agency's appeal to the panel's track record by the rate at which panel membership has changed over time. Such a practice would reduce, at the margin, agency incentives to manipulate the composition of expert panels by substituting new members whose viewpoints will predictably track the agency's own.[111]

---

110. *Cf.* Matthew C. Stephenson, *Bureaucratic Decision Costs and Endogenous Agency Expertise*, 23 J.L. ECON. & ORG. 469, 472 (2007).

111. In the 1980s, EPA was accused of manipulating the membership of its Science Advisory Board in this way. A Reagan administration "hit list" was discovered containing the names of advisors whose views were no longer sympathetic to that of the administration. Most of the

By doing so, the agency would also be undercutting its own ability to appeal to the panel's track record as a basis for departing from its views.

I. Comparative Qualifications

In some cases, agencies will have valid reason to reject nose counting—to decline to follow a (super)majority view of the panel—because the agency validly weights votes by the qualifications of the voters. In the black lung cases it is perfectly rational for the agency adjudicator to believe the diagnoses of two doctors with special qualifications in detecting pneumoconiosis over the diagnoses of three general practitioners. Counting weighted votes is an epistemic improvement, so long as the weights track competence and—a crucial qualification—so long as the discounted voters bring no cognitive diversity to the group. If the latter condition does not hold because the discounted voters bring new perspectives or have views that, by training or profession, are likely to be uncorrelated with the views of highly competent experts, then the logic of the Jury Theorem suggests that the group's overall epistemic performance will be better than the views of even its most expert members.[112]

J. Bias

In other cases, agencies might be able to point to systematic bias among members of the panel in order to impeach the epistemic warrant for their conclusions. If statutes create a panel composed, for example, largely of experts drawn from a certain discipline, profession, or industry, and if the panel's recommendations track apparent disciplinary or professional biases or industry interests, there is valid ground for concern. In *National Nutritional Foods Ass'n v. Califano*,[113] the Second Circuit heard a challenge to the composition of an advisory committee charged with making recommendations about whether consumer warnings should be placed on packaged dietary supplements. Although the challenge was dismissed on procedural grounds not relevant here, the court noted that the panel was composed solely of physicians, that the physicians were

scientists on the list were, indeed, "retired" by the Reagan EPA. Nicholas A. Ashford, *Advisory Committees in OSHA and EPA: Their Use in Regulatory Decisionmaking*, 9 SCI. TECH. & HUM. VALUES 72, 72 n.1, 77 n.16 (1984).

112. Ladha, *supra* note 77, at 619.

113. Nat'l Nutritional Foods Ass'n v. Califano, 603 F.2d 327 (2d Cir. 1979).

"understandably leaning in favor of medical supervision of [the underlying substances],"[114] and that the panel's skewed composition "directly implicates the concern Congress addressed in [the Federal Advisory Committee Act], that agency action might be dominated by one particular viewpoint."[115] By analogy, the same concerns suggest that agencies will have better second-order reason to depart from panel recommendations when the panel's composition is narrowly defined or drawn predominantly from industry.

Here there are two relevant interpretations of the vague term "bias": motivational and epistemic. In the former sense, bias means that the expert is not even trying to reach the right answer, as opposed to the answer that benefits the expert's firm or career; in the latter sense, bias means that the expert has blind spots arising precisely from specialized training or knowledge. In Jury Theorem terms, bias may thus be understood either as a simple violation of the condition of sincere voting, at the level of individuals, or else as a concern about the positive correlation of errors across the group. If members of a scientific subdiscipline, profession, or industry have highly correlated perspectives, it is less likely that errors will wash out at the group level.

In many cases, however, the underlying statutes require a mix of professions, disciplines, and perspectives, specifying with particularity how the panel should be composed. Under the CAA, for example, CASAC comprises "at least one member of the National Academy of Sciences, one physician, and one person representing State air pollution control agencies."[116] Likewise, the ACCV is comprised of health experts, members of the general public (two of whom have children who have suffered vaccine-related injury or death), lawyers, and officials from relevant agencies. Even more pointedly, the National Institutes of Health (NIH) Revitalization Act of 1993[117] authorizes the secretary of HHS to appoint an Ethics Advisory Board.[118] If such a board is appointed, it must have between fourteen and twenty members, including at least one attorney, one ethicist, one practicing physician, one theologian, and no fewer than one-third but

114. *Id.* at 334.

115. *Id.* at 336.

116. 42 U.S.C. § 7409(d)(2)(A) (2006).

117. National Institutes of Health Revitalization Act of 1993, Pub. L. No. 103-43, 107 Stat. 126 (codified as amended in scattered sections of 42 U.S.C.).

118. 42 U.S.C. § 289a-1(b)(5)(2006).

no more than one-half of the members may be "scientists with substantial accomplishments in biomedical or behavioral research."[119] Most generally, the Federal Advisory Committee Act indirectly requires that "the membership of . . . advisory committee[s] . . . be fairly balanced."[120]

Requirements of this sort trade off some scientific competence, at the margin, for greater representation of affected interests and reduced correlation of errors at the group level. Individual epistemic competence of the panel members is just one good, which should be optimized, not maximized; balanced panels of this sort can create overall gains by sacrificing some expertise for a reduced chance that the biases of any one affected interest will dominate. In the case of the NIH Revitalization Act of 1993, the lower bound on biomedical scientists (at least one-third of the panel) promotes expertise, and the upper bound (no more than one-half of the panel) promotes epistemic diversity by restricting the representation of a particular scientific subfield. Similarly, a diversity of affected interests minimizes the chance that panel members will deliberately falsify an appearance of consensus in order to maximize their influence. Because diverse interests will predictably have crosscutting agendas, the chances that the panel can agree on a single position are reduced; more likely is open disagreement, providing more information to agencies and reviewing courts.

The cost of epistemic diversity is a slightly increased chance that some members of a panel will believe something truly bizarre—that lead is beneficial or that last week's cold weather shows that climate change is not occurring.[121] Under the approach I am suggesting, the agency would be barred from simply rejecting these conclusions on the first-order ground that the agency knows them to be false. Yet in order for these bizarre views to make any difference, they must (1) obtain the agreement of a majority of panel members, expert as well as nonexpert, (2) under circumstances in which the agency has no valid second-order reason to reject the panel's conclusions. The wackier the error, the less likely it is that those additional two conditions will hold, so this is another self-limiting problem.

The consequences for administrative law and judicial review are straightforward. When requirements of epistemic diversity are in

119. *Id.* § 289a-1(b)(5)(C).

120. 5 U.S.C. app. 2 §5(b)(2) (2006).

121. Thanks to Lisa Heinzerling for these examples.

place, courts should be reluctant to accept an agency's appeal to systematic disciplinary, professional, or industry-based bias. When an expert panel is drawn solely from a narrow professional subcategory or subdiscipline, or staffed largely by representatives from a particular industry or segment of industry, an agency appeal to bias should be taken more seriously.

To illustrate the right analysis, CASAC is a possible target for an agency challenge of this sort; but the challenge is weak and should fail. The overall Science Advisory Board is required by the board's charter to be comprised of "independent experts in the fields of science, engineering, economics, and other social sciences to provide a range of expertise required to assess the scientific and technical aspects of environmental issues."[122] Here, all panel members are scientists and technicians, but not from any particular discipline or scientific field, in accordance with the statute's mandate for "diversity" and "a range of expertise." CASAC, a committee of the Science Advisory Board with an independent statutory charter, is diverse across professions as well as within scientific disciplines; it contains, at a minimum, a physician and a state environmental official, in addition to scientists and engineers of various types.

K. Groupthink and Judgment Falsification

In general, agencies should be allowed to depart from the views of experts when there is good reason to think that herding or groupthink has occurred. In these cases, the agency has a valid second-order reason for discounting an expert majority view, or consensus; if the problem is serious, the agency's first-order view may be the only independently formed first-order view in the field. Here there is a good deal of overlap with the vague idea of "bias," interpreted in its epistemic rather than motivational sense. Even if the goal of all panel members is to get the answer right, copying of others' views, although individually rational, can make the group decision uninformed.

Although various forms of groupthink are a real concern, there are two relevant cautions. First, the mere fact that some experts on a panel follow the views of other experts does not amount to groupthink, or necessarily reduce the overall epistemic competence of

122. U.S. EPA, U.S. EPA CHARTER: EPA SCIENTIFIC ADVISORY BOARD ¶ 10 (2007), *available at* http://yosemite.epa.gov/sab/sabproduct.nsf/WebBOARD/currentcharter?OpenDocument.

the group. If panel members defer to a highly competent opinion leader, group epistemic competence can increase overall. Equivalently, herding or information cascades cannot be inferred from the bare fact that some members of the panel copy the views of others. If the copying members have *metaexpertise*—if they are quite adept at figuring out who among them are the best experts—then copying can actually improve the group's overall performance.[123] The copiers "may be poor meteorologists, but good judges of meteorologists."[124]

Second, the risk of groupthink is sensitive to the composition and structure of the panel and the decisionmaking processes it uses. Accordingly, agencies will have valid second-order reasons to discount the consensus of expert panels when they can point to features of the panel's composition, structure, or decisionmaking process that raise red flags. When an expert panel is all drawn from the same subfield; when experts vote or express judgments in sequence and with knowledge of each others' views; when experts are uncompensated or when their compensation is a function of individual rather than group performance; and when a panel's views are completely unanimous on an issue the agency has independent reason to think is at least difficult, then the risk of harmful groupthink is at an apogee and reviewing courts should take the agency's second-order concerns most seriously.

L. The Unanimity Dilemma

Special problems arise when an expert panel is unanimous or reports consensus without a formal vote. In these cases, both agencies and reviewing courts are in something of an epistemic dilemma. Unanimity can arise either because all experts on the panel, whatever their biases, are receiving a strong and uniform signal from reality about an issue of fact or causation. It can also be a sign of herding, cascades, or judgment falsification. Unanimity is epistemically ambiguous. By contrast, the most powerful expert consensus is a supermajority that is not unanimous. The open dissent shows that the supermajority's view has been epistemically tested by vigorous disagreement, yet has still prevailed.

123. David Coady, *When Experts Disagree*, 3 EPISTEME 68, 71 (2006).

124. *Id.* at 72.

Unanimity, then, is both potentially powerful and potentially suspect. Under what conditions is it most likely to be one or the other? When reviewing courts have a prior belief that the issue is an easy one, or that the facts lean strongly in one direction, unanimity is best taken as a warning that an agency's contrary view would lack any basis in fact. In most regulatory issues that reach the stage of final agency action, however, and for which an expert panel has been appointed, it is unlikely in the extreme that the issue is antecedently known to be an easy one, or that the facts overwhelmingly favor one view. In these circumstances, unanimity is likely to be suspect, whereas disagreement within the panel should actually increase the agency's, and the reviewing court's, epistemic confidence in the conclusions of the panel majority or supermajority.

One implication, somewhat counterintuitive, is that agencies should not generally be allowed to impeach the conclusions of a panel majority by claiming that disagreement within the panel shows that the minority view is reasonable. It may be so, but the question is where agencies and reviewing courts should place their epistemic bets: with the minority or with the majority. The logic of expert aggregation suggests that placing epistemic bets on panel majorities is the better course, on average. And if panels collectively desire to maximize the chance the agency will adopt their recommendations, a legal rule that allows agencies to impeach panel conclusions by pointing to reasonable dissent will give panels incentives to falsify an appearance of consensus, thus suppressing useful information.

In the 2006 controversy over fine particulate matter, the administrator pointed out that the CASAC subcommittee was not unanimous, as a basis for differing from its recommendations. This argument is doubly misguided. The questions at issue were hardly simple, whatever their correct resolution. CASAC's decision, endorsed by a large but nonunanimous supermajority, actually offered firmer grounds for epistemic confidence than a unanimous one, by showing that relevant arguments had been ventilated; using non-unanimity as a basis for rejecting the panel's views gives CASAC members a heightened incentive to create a false appearance of consensus in the future. Happily, the reviewing court paid no heed to the administrator's observation.[125]

125. *See* Am. Farm Bureau Fed'n v. EPA, No. 06-1410, 2009 WL 437050 (D.C. Cir. Feb. 24, 2009) (per curiam) (failing to discuss the Administrator's attempt to impeach CASAC's recommendation).

M. An Undemanding Presumption

Although we have canvassed a range of second-order reasons to which agencies might appeal, the judicial task under this framework is not terribly demanding. The structure of the inquiry is a simple presumption: unless agencies can clearly establish that a valid second-order reason exists, the agency is obliged to adopt the expert panel's factual findings. Although the second-order reasons set out above will be unfamiliar to some judges, the concepts are not difficult in themselves. As we have seen, FDA recently drew upon the academic literature on information cascades to formulate voting rules for its advisory panels,[126] so judges should have little difficulty with these tools. In comparative terms, the suggested inquiry is no more demanding, and quite possibly less demanding, than standard hard look review of the agency's first-order reasons, which, effectively, requires judges to decide whether the agency's substantive scientific and technical claims are minimally plausible.

CONCLUSION

I will conclude by returning to the puzzles laid out at the beginning of the introduction and indicating how they should be resolved under the framework I have suggested. In reverse order:

• When an expert panel is evenly split, as in *Oregon Natural Resources Council v. Daley*, the agency should have a tiebreaking vote, and a first-order reason should be sufficient. The agency's view counts for at least one, although for no more than one. On jury-theoretic grounds, the views of half of the panel plus the agency's first-order views are more likely to be correct than the views of the other half of the panel. More pragmatically, there is no obvious alternative to letting the agency break the tie.

• In the black lung cases, and other cases in which a crisply defined issue of fact or causation is presented to an agency adjudicator, nose counting should generally be an acceptable basis for decision under the arbitrary and capricious or substantial evidence tests. Nose counting is a rational second-order epistemic strategy; to require an administrative law judge to disgorge a first-order reason, such as a medical opinion, in order to arbitrate an expert disagreement guarantees arbitrary decisionmaking.

126. U.S. DEP'T OF HEALTH & HUMAN SERVS., *supra* note 79, at 5.

• Agencies should be obliged to follow the (super)majority view of an expert panel, as to matters of fact, causation, or prediction, unless they can give a valid second-order reason for rejecting that view—for example, that the panel's composition made it inadequately diverse, or that its decisionmaking processes or structure made groupthink or judgment falsification a serious risk. A hard case arises where an expert panel is unanimous or reports consensus. Here, the decision represents either of two epistemic extremes: maximally persuasive or epistemically suspect.

• EPA's decisions, in 1979 and 1997, to follow the CASAC (super)majority were both correct. Neither was consistent with its 2006 decision, in a rulemaking on the same questions, not to follow the CASAC supermajority; thus the D.C. Circuit correctly held that the 2006 decision was inadequately reasoned. When, as in the 2006 rulemaking, a large supermajority of a nonunanimous panel reaches a certain conclusion, the agency will rarely have a valid second-order reason to depart from the panel's view. Because of the presence of dissenters, the epistemic quality of the expert majority is at an apogee. The administrator's observation in the 2006 rulemaking that the expert panel was not unanimous actually undermined his decision to depart from its conclusions; it gave more reason, not less reason, to consider his views erroneous.

My largest point, running throughout the foregoing, is that administrative law inconsistently aggregates expert opinions. Some decisions are unduly and even irrationally reluctant to grant a formal role to nose counting of experts or expert panels. Voting by expert panels is likely, on average, to be an epistemically superior mechanism for determining facts and causation, and for making predictions, than is the first-order judgment of agency heads in rulemaking or adjudication. Many believe that agencies should be politically accountable to Congress; there are equally strong grounds to think that they should be epistemically accountable to the parliament of the experts.