Mapping Scientific Agreement and Disagreement: The Case for a Dedicated Center

Haixin Dang[1], Jesse Hamilton[2], Rory Jubber[3], Owen Towler[3], and Peter Vickers[4]

[1] Department of Philosophy, Case Western Reserve University, Cleveland, United States of America.

[2] Blavatnik School of Government, University of Oxford, Oxford, UK

[3] Department of Science and Technology Studies, University College London, London, UK

[4] Department of Philosophy, University of Durham, Durham, UK

All authors contributed equally to this article.

Corresponding author: Peter Vickers (peter.vickers@durham.ac.uk)

Mapping Scientific Agreement and Disagreement: The Case for a Dedicated Center

Abstract

From 2022 to 2025, initial steps were taken toward establishing a dedicated center for systematically mapping scientific agreement and disagreement. This paper examines whether creating such a body is both desirable and feasible. We summarize lessons from three pilot studies (COVID-19; the existence of extraterrestrial life; claims of extraterrestrial life discovery), outlining key design choices. We then catalog and analyze prominent objections, drawing on perspectives from philosophy of science, social epistemology, STS, and political philosophy. We distinguish the proposed remit of such a center from Collins and Evans's Owls: unlike the Owls it would be non-statutory, non-directive, and descriptive, providing auditable snapshots and longitudinal baselines of expert judgment rather than policy or governance prescriptions. Our conclusion is that, under appropriate design constraints, a center of this kind is operationally tractable and normatively

defensible, and, if properly designed, can illuminate patterns of disagreement, uncertainty, and agreement without licensing technocracy or elitism.

1. Introduction

It is sometimes insisted that scientific consensus is epistemically important (cf. Oreskes 2019). A split scientific community indicates that a matter is not settled, and debate continues. The split may be 50:50, 60:40, 70:30, or perhaps even 80:20. But an absolutely solid consensus is supposed to be of special significance. In the early months of the COVID-19 pandemic, the public were told, in many countries, that there was no consensus on the efficacy of face masks, apparently justifying an individual's personal choice at that time (cf. Greenhalgh et al. 2020). By contrast, in the case of human-caused global warming, scholars have gone to great lengths to emphasize the strength of agreement among experts—not only that climate change is real, but that it is caused by human activities. In such cases, we are encouraged to treat consensus as an important signal of evidential strength.

Framed in this way, debate tends to focus on a single question: whether consensus exists. But this focus is overly narrow. Consensus is best understood as a limiting case of a more general phenomenon: the distribution of expert judgment within a scientific community. At any given time, domain-relevant experts assign differing levels of confidence, or credence, to a scientific claim. Taken together, these credences form a distribution. Consensus arises when that distribution converges sufficiently; in many cases, however, such convergence does not occur. What matters, then, is not simply whether consensus is present, but how expert judgment is structured across the community.

This broader perspective preserves features that are otherwise obscured. A community may exhibit tight clustering around a moderate level of confidence, broad dispersion across the full range of possibilities, or distinct subgroups with competing views. These patterns—variance, clustering, and persistent disagreement—are themselves epistemically significant.

This perspective is particularly important from the standpoint of non-experts. Faced with the question “What should I believe?”, a layperson has limited options. Conducting one’s

own research is usually inadvisable in domains requiring specialized expertise (Levy 2022, Ch. 4). Consulting a single expert may be misleading if that expert occupies a minority position within their field, or even a relatively weak majority position. Media coverage, likewise, typically presents only a small and potentially unrepresentative sample of views. What is missing, in many cases, is access to how expert judgment is distributed across the relevant scientific community: not just what one or two individual experts think, but how views are patterned across the field as a whole.

Consider, for example, recent claims concerning possible biosignatures in the atmosphere of exoplanet K2-18 b. At the time of writing, no consensus has formed. Experts differ in their assessment of both the evidential basis and the interpretation of the data. In such cases, the absence of consensus does not imply epistemic indeterminacy. What matters is how expert judgment is distributed: whether it is tightly clustered at low confidence, broadly dispersed, or divided into competing camps. These features inform how a non-expert's confidence ought to be calibrated under conditions of ongoing scientific uncertainty.

On this view, it is striking that no standing body currently exists tasked with systematically mapping scientific community opinion in this broader sense. Even in domains where agreement is frequently discussed, such as climate change, attempts to quantify expert views have been ad hoc and methodologically heterogeneous. The widely cited figures of 97%, 100%, >99%, and 98.7%—associated with Cook et al. (2016), Powell (2019), Lynas et al. (2021), and Myers et al. (2021), respectively—are the products of small teams employing differing methods. They capture particular snapshots, but there is no persistent infrastructure for tracking how expert judgment is distributed, how it evolves over time, or how it varies across communities.

Given that consensus, or its absence, is widely taken to have epistemic significance, why are there no formal bodies or mechanisms that attempt to track scientific community opinion in a systematic way? One possible explanation lies in a familiar skeptical narrative. On this view, emphasis on consensus is misguided or even dangerous. Drawing on a broadly Popperian image of science as a process of conjecture and refutation, critics warn that consensus can entrench false views, suppress dissent, and encourage conformity. Scientists,

it is said, must be free to pursue minority and heterodox positions. This perspective is often illustrated with well-known slogans: “[S]cientific truth is not established by consensus”, “The work of science has nothing whatever to do with consensus”, and “Scientific truth is not decided by majority vote” (Vickers 2023; Jorm 2025). A related sentiment holds that “Science is a matter of evidence, not what a majority of scientists think” (Jorm 2025), and similar claims continue to appear in contemporary scientific discourse (Thorp 2025).

A quite different set of concerns arises from the sociology and politics of science. From this perspective, initiatives that purport to identify and emphasize ‘consensus’ are viewed as politically charged. Measuring agreement may reinforce existing hierarchies, marginalize dissenting communities, or be co-opted by institutional actors to define what counts as legitimate knowledge. The concern here is less about epistemic error than about power: who gets to define orthodoxy, and whose voices are excluded.

The approach developed in this paper seeks to move beyond these opposing narratives centered on ‘consensus’. Rather than treating consensus as either epistemically decisive or epistemically suspect, we focus on the systematic mapping of

expert judgment distributions. On this approach, consensus is neither the sole object of interest nor something to be dismissed, but one possible configuration within a wider evidential landscape.

The initial steps toward establishing the Centre for Scientific Community Opinion Polling and Evaluation (C-SCOPE) during 2022–2025 have provided a proof of concept for this approach. These efforts have also generated a fresh articulation of concerns, some familiar and some novel. In this paper, we collect and analyze the principal objections, drawing on perspectives from philosophy of science, social epistemology, STS, and political philosophy. The author team reflects this diversity of backgrounds, and we aim to address the concerns from multiple disciplinary viewpoints.

As a convenient reference point, we here list the concerns that will be tackled in sections 3, 4, and 5 of this paper:

EPISTEMOLOGICAL CONCERNS

- 3.1. Science doesn't work by a show of hands.

- 3.2. Surveying whole communities doesn’t tell us very much.

- 3.3. Measuring or publicizing consensus promotes conformity.

COMMUNICATION CONCERNS

- 4.1. The data produced could result in lowering public trust in science

- 4.2. This is a form of the much maligned ‘deficit model’.

POLITICAL CONCERNS

- 5.1. The proposed concept bears a striking resemblance to Harry Collins and Robert Evans’ (Collins and Evans 2017) idea of the Owls, inheriting all its problems as a result.

- 5.2. The power held by such a standing body would be technocratic and elitist.

In Section 2 we explain what was done across 2022-25, including the new methodology proposed, the data collected, and how the idea could feasibly be developed in the near future. In Sections 3-5 we take the listed concerns in turn. In Section 6 we take stock and conclude. Notwithstanding the list of concerns, we find that the kind of standing body envisioned is both operationally tractable and normatively defensible.

2. The ‘C-SCOPE’ concept and activities to-date

Initially termed the *Institute for Ascertaining Scientific Consensus* (IASC), and subsequently rebranded as C-SCOPE (Centre for Scientific Community Opinion Polling and Evaluation), the idea is to establish a standing, non-statutory body dedicated to systematically mapping scientific community opinion.[1]

[1] The explicit reference to ‘consensus’ in the original name ‘IASC’ was subsequently judged to be misleading. In practice, the most informative findings to date concern disagreement, uncertainty, and the structure of scientific opinion (and opinion change) rather than consensus per se. Alternative names considered during the rebranding process included the

At its core, C-SCOPE is descriptive rather than prescriptive. Its function is not to adjudicate truth or issue policy recommendations, but to generate transparent, auditable snapshots of scientific community judgment, including longitudinal baselines tracking how opinion evolves over time. In this respect, C-SCOPE is best understood as contributing to the epistemic infrastructure of science and public discourse: it provides structured information about what scientific communities think, without attempting to determine what should be done.

2.1 Conceptual design

The approach rests on three core design principles.

First, surveys are ultra-brief, typically taking only seconds to complete. This minimizes respondent burden (survey fatigue) and makes it possible to conduct repeated measurements

Centre for Mapping Scientific Agreement and ISOBAR (Institute for Scientific Opinion Barometry), before the Steering Committee selected C-SCOPE via a formal vote. In May 2026, C-SCOPE was formally established as a research center at Durham University.

over time, thereby enabling the construction of longitudinal datasets tracking the evolution of scientific opinion.[2]

Second, responses are anonymized, with each participant receiving a unique, single-use token. This preserves anonymity while allowing limited categorization by broad metadata (e.g. field and institution), striking a balance between analytical utility and privacy.

Third, survey invitations are sent *either* via institutional representatives ('spoke reps') at each spoke in the network, or else are sent via the PI with a trusted figure in CC. Either option reduces the risk of surveys being treated as unsolicited external requests, and thus improves participation rates.

These principles reflect a deliberate methodological trade-off. Traditional surveys prioritize depth, collecting extensive demographic and attitudinal data at the cost of poor response rate, high opt-out rate, and limited possibility of repeats or follow-ups. C-SCOPE prioritizes breadth, frequency, and endurance, enabling large-scale, repeatable measurement of community-level opinion.

[2] Survey-fatigue is very real and has been explored in various studies. See for example Lavrakas (2008), De Koning et al. (2021), and Brown et al. (2024).

2.2 Pilot studies (2022–2025)

Three pilot studies demonstrate the operational viability and epistemic potential of this approach.

(i) COVID-19 proof-of-concept (2023)

The first pilot established a hub-and-spoke network across 30 universities in 12 countries, coordinated from Durham University. A single-statement survey concerning the viral cause of COVID-19 was distributed internally via institutional representatives or ‘spoke reps’ (Vickers et al. 2024).

A total of 20,085 scientists were invited and 6,807 responded, yielding a response rate of 33.9%, with rates exceeding 55% at several institutions. This is unusually high for surveys of scientific opinion and demonstrates the effectiveness of ultra-brief surveys combined with trusted distribution channels.[3]

This pilot demonstrated the feasibility of:

- large-scale international coordination;
- token-based anonymous participation;

[3] The IRIS survey on research integrity (Allum et al. 2023) had a response rate of 7%, for example.

- high response rates in scientific populations;
- rapid data collection.

(ii) Astrobiology surveys (2024)

The second pilot applied a refined methodology to the global astrobiology community, alongside a comparison group of non-astrobiologists. Invitations were issued centrally, with a recognized figure in the field in CC.

Three statements probed views on the likely existence of basic, complex, and intelligent extraterrestrial life. Results revealed a clear gradient:

- Basic life: ~86–88% agreement;
- Complex life: 67.4% agreement;
- Intelligent life: 58.2% agreement.

This pattern is significant. It shows that C-SCOPE does not merely identify the presence or absence of consensus, but can reveal structured gradients of agreement, reflecting how scientific communities calibrate belief under uncertainty. A range of other notable results emerged. For example, multiple indicators confirmed that the agreement-drop from basic to

complex life was significantly larger than the drop from complex to intelligent life (see Vickers et al. 2025).

(iii) Rapid-response surveys on extraterrestrial life detection claims (2025)

The third pilot extended the approach to real-time scientific developments, focusing on two high-profile claims of possible extraterrestrial life detection: atmospheric signatures on exoplanet K2-18 b, and the Cheyava Falls rock on Mars (see Vickers et al. 2026).

Results showed:

- Low agreement for K2-18 b (6.6%) with substantial disagreement;
- Higher but still cautious agreement for Cheyava Falls (15.1%);
- Significant shifts in neutrality and disagreement distributions.

These findings demonstrate that C-SCOPE can capture early-stage community reactions, quantify uncertainty, and establish comparative baselines across cases. Crucially, they reinforce that the project is not primarily concerned with

detecting consensus, but with mapping the full distribution of scientific opinion.

2.3 Network expansion and methodological consolidation

Across these pilots, the network expanded from 30 to 90 institutions, across 34 countries, each with a named representative. A beta-testing group of over 100 scientists was established, and studies were preregistered using the Open Science Framework.[4]

These developments embed transparency, replicability, and auditability into the model. Because sampling methods, survey platform specifications, and outputs are public (Mitchell Finnigan et al. 2023), independent teams could in principle reconstruct comparable surveys, providing an integrity check on the results.

2.4 Value-ladenness and statement selection

[4] Pilot studies were already preregistered, using the Open Science Framework: see osf.io/r4sy2 and osf.io/xusv7.

A central methodological issue concerns the selection and formulation of survey statements, and in particular the extent to which such statements are value-laden. Some candidate statements are heavily value-laden, in that their interpretation depends substantially on normative judgments rather than empirical evidence alone. For example:

“This vaccine is good.”

Here the evaluative content is explicit. Respondents must draw on normative judgments concerning benefits, harms, and broader social or ethical considerations. The statement does not admit of a purely evidential interpretation.

Other statements are less overtly evaluative, but still significantly value-laden. Consider:

“This vaccine is safe.”

The concept of ‘safety’ is inherently threshold-dependent. Respondents may interpret it in terms of acceptable inductive risk, population-level versus individual harms, or broader cost–benefit considerations. These are not purely scientific questions, but involve implicit value judgments about what levels of risk are tolerable and for whom. Moreover, such statements are often embedded within ongoing political

controversies, raising further concerns about how responses are formed and interpreted.

By contrast, consider a statement such as:

“COVID vaccines elicit an immune response in humans.”

Here the interpretation is much more tightly constrained by empirical considerations. Although no statement is entirely free of background assumptions, responses are driven primarily by evidential and inferential judgment rather than contested normative thresholds.

Statements of the first two kinds are methodologically problematic within the C-SCOPE framework. They risk conflating scientific judgment with normative evaluation, thereby obscuring what is being measured. The third type, by contrast, more cleanly targets expert judgment about the evidential status of a claim.

A similar contrast can be seen in the second pilot study:

“It is likely that extraterrestrial organisms with advanced cognitive abilities comparable to or superior to those of humans exist somewhere in the universe.” (see **Fig. 1**).

[FIGURE 1 HERE]

This statement is not value-free: it involves probabilistic judgment and interpretation of uncertain evidence. However, it is substantially less value-laden than the first two vaccine examples above. Its interpretation depends primarily on scientific reasoning under uncertainty, rather than on contested normative thresholds, and it is comparatively less entangled with contested social and political values.

The contrast illustrates an important principle. C-SCOPE does not assume that there exist ‘pure’ scientific statements that are entirely free of values. Rather, statements vary in the extent to which their interpretation is shaped by normative commitments. At one end are statements whose meaning is dominated by evaluative considerations; at the other are statements whose interpretation is predominantly evidential and inferential.

The aim, therefore, is not to eliminate value-ladenness altogether, but to avoid statements where it is dominant or distorting. In practice, this involves selecting statements whose interpretation is sufficiently constrained by shared evidential standards that responses can be meaningfully

interpreted as reflecting expert judgment about the state of the evidence.

This constraint should not be confused with a sufficient criterion for statement selection. A statement may be comparatively free from problematic value-ladenness while nevertheless being too trivial or uninformative to warrant surveying. The examples above are intended to illustrate differences in normative interpretation, rather than to identify statements that C-SCOPE would select. Statement selection must therefore consider not only whether a claim is sufficiently constrained by evidential considerations, but also whether mapping expert judgment concerning it would likely illuminate a scientifically significant question.

2.5 Statement selection and validation procedure

C-SCOPE adopts a multi-stage procedure for statement selection and validation. The aim is to ensure that statements are clear, methodologically appropriate, and robust under variation in interpretation.

In outline, the procedure involves:

1. Initial formulation by a Working Group, which drafts candidate statements and provides a brief justification.
2. First-stage review by a Steering Committee, assessing clarity, value-ladenness, substantive informativeness, and potential sources of bias, with the option to reject or request revision.
3. Beta-testing across a distributed group of scientists, inviting feedback.
4. Revisions by the Working Group in light of this feedback, accompanied by a brief record of how concerns have been addressed.
5. Final approval by the Steering Committee prior to preregistration.[5]

Judgments of substantive informativeness are necessarily context-sensitive and cannot be reduced to a mechanical selection rule.

This process combines content expertise, methodological oversight, and distributed interpretive testing. It provides a transparent and auditable mechanism for filtering out

[5] In some cases, translation of the resultant statement(s) would be needed - as well as translation of the survey-invitation email. This requires extra steps in the validation procedure.

problematic statements and refining those that are retained. In doing so, it directly addresses concerns about bias, framing, and legitimacy, which are examined in greater detail in later sections.

2.6 Summary and forward trajectory

Taken together, the pilot studies and methodological developments demonstrate that C-SCOPE is:

- operationally tractable, capable of generating large, international samples with high response rates;
- methodologically distinctive, prioritizing repeatability and scale;
- empirically informative, revealing structured patterns of agreement and disagreement.

Most importantly, the project establishes the feasibility of a persistent measurement infrastructure for scientific community opinion. Despite the central role often attributed to consensus in both scientific and public discourse, there currently exists no systematic mechanism for tracking patterns of agreement and disagreement within scientific

communities on specific claims, hypotheses, and interpretive judgments. C-SCOPE begins to address this gap.

The eventual scope of such an infrastructure should not be assumed in advance. The pilot studies already demonstrate sufficient epistemic usefulness to make systematic exploration of wider applications worthwhile. The astrobiology studies, in particular, show that community-level polling can reveal scientifically interesting patterns of agreement, disagreement, and uncertainty even in a domain with relatively little immediate policy relevance. How far the methodology can be extended to other domains, including more interdisciplinary, politically salient, or policy-relevant questions, is an empirical and methodological question for further work. Establishing these scope conditions is itself part of the rationale for developing C-SCOPE as a standing research infrastructure.

The remaining sections assess the principal objections to such an enterprise.

3. Epistemological concerns

Part of the motivation for rebranding to C-SCOPE from the Institute for Ascertaining Scientific Consensus was to avoid the connotation that it is a body whose end goal is *achieving* consensus data rather than *revealing* patterns of agreement and disagreement within an expert community. Consensus is only one special case of community agreement, and it is something we are interested in measuring alongside many other distributions of expert judgment. Nevertheless, efforts to measure scientific consensus have prompted a range of conceptual and epistemological objections. Critics question whether consensus has genuine epistemic significance, whether quantifying it risks homogenizing scientific thought, and whether such projects rely on contentious philosophical assumptions. These concerns raise important issues about how scientific communities generate, coordinate, and communicate expert judgment. In what follows, we examine the most prominent of these objections, assess their historical and philosophical basis, and consider what they imply for the value and limitations of measuring agreement within scientific fields more broadly.

3.1 Science doesn't work by a show of hands.

The very idea of a ‘scientific consensus’ sometimes elicits reflexive dismissals. Critics counter with a familiar litany of responses:

- “[S]cientific truth is not established by consensus.” (Bergin and Strupp 1972, 145)
- “Let’s be clear: The work of science has nothing whatever to do with consensus… […]. In addition, let me remind you that the track record of consensus is nothing to be proud of.” (Michael Crichton 2003, lecture at Caltech; see Jorm 2025, 5)
- “Science is a matter of evidence, not what a majority of scientists think.” (John Kay writing in the Financial Times, 2007; see Jorm 2025, 6)
- “Scientific truth is not decided by majority vote, of course (ask Galileo).” (Sharp 2009, 206)
- “Scientific truth is not determined by consensus. In science the majority can be, and often have been, wrong. There is a great deal of peer pressure in the

science community, which can stifle objections to a popular theory.” (Elsdon-Baker 2009, 215)

- “When everyone believed that the world was flat, they were wrong. The truth is not decided by a show of hands.” (Curnow 2012, Chapter 4)

The list of quotes could be extended ad nauseam. Even more recently, upon learning of the PLOS ONE study (Vickers et al. 2024), Sabine Hossenfelder cursorily dismissed the project on her popular YouTube channel as a ‘horrible idea’ (Hossenfelder 2024), arguing that it is not the opinions of scientists that matter but rather the scientific evidence itself.

All these reactions share a common logic. Scientific consensus —as a social phenomenon—is considered extraneous to the real scientific process. They treat the idea of scientific consensus as if its defenders were making the naive claim that truth depends on a vote, or that scientific knowledge is governed by majority rule. Frequently, these objections also gesture toward a Popperian picture of science as advancing solely through bold conjectures and decisive refutations, as if any appeal to consensus were inherently antithetical to scientific progress.

In a similar vein, there is an underlying suspicion that groups can be irrational or unreliable. Experts may engage in groupthink—conforming their opinions to their peers during deliberation—rather than assessing the evidence directly (Solomon 2006). Consensus is then rendered pathological rather than epistemically valuable.

In challenging the very idea of scientific consensus, these objections often sideline any serious analysis of collective expert judgment in science. Little attention is given to how collective expert judgments actually operate within scientific communities or the kinds of epistemic information that different distributions of agreement and disagreement might convey. Yet these patterns of opinion may reveal where evidence is regarded as robust, where uncertainty remains substantial, how methodological or disciplinary divisions shape evaluation, and whether a field is moving toward convergence or fragmentation. The distribution of expert judgments within a scientific community is therefore itself an important source of insight into how science functions and should not be dismissed.

It is also worth recognizing that some attacks on consensus are not purely philosophical. Michael Crichton’s now-famous

anti-consensus remarks, for example, appeared in a lecture that cast doubt on research into anthropogenic climate change. In debates over climate change and the health effects of smoking, a few dissenting voices have been amplified to create the impression of substantial scientific disagreement and sow public doubt—to the detriment of public policy (Oreskes and Conway 2010). De Melo-Martín and Intemann (2018, 3–4, 20–21) help clarify why this is troubling without treating all scientific disagreements as themselves pathological. They argue that expert disagreement can expose errors and strengthen the justification for consensus, even when the dissenting view is mistaken.[6] But they also examine what they call "normatively inappropriate dissent" (NID): dissent that fails to serve these epistemic purposes and can cause wider harms. They caution that distinguishing NID from epistemically productive dissent can be extremely difficult. In the climate and tobacco cases, the concern is the selective amplification of dissent to misrepresent the judgment of the

[6] Furthermore, Dellsén (2025) argues that, all else being equal, a strong consensus with a small amount of expert dissent may give non-experts better reason to trust a claim than complete unanimity. The presence of dissent makes it less likely that widespread agreement is merely the result of conformity. On this account, dissent can strengthen the evidential significance of a consensus rather than undermine it.

wider expert community. These attacks on the scientific consensus are not conducted in good faith; rather, they are meant to undermine the credibility of entire research programs.

Furthermore, dismissals of scientific consensus do not show that consensus actually lacks *epistemic* significance; rather, they reveal a misunderstanding of what its significance could be. Critics conflate the logical point that truth is not constituted by a show of hands with the epistemic point that collective expert judgment can help us assess the weight of evidence. Scientific evidence does not speak for itself: it must be interpreted, evaluated, and situated within a broader body of inquiry. In technically complex or highly specialized fields, this work depends on expert judgment.

Philosophers of science have long recognized that scientific consensus can carry epistemic weight. In 2004, Oreskes argued that the peer-reviewed climate literature reflected a consensus on anthropogenic climate change, despite public claims of substantial expert disagreement (Oreskes 2004). In her later account, she explains why such community agreement deserves our trust. Drawing on Longino's (1990) notion of "transformative interrogation," Oreskes argues that

sustained criticism within a diverse scientific community can expose errors and biases. A “critically achieved consensus” may therefore be the best available basis for trusting a scientific conclusion, though it cannot guarantee its truth (Oreskes 2019, 51–55, 59). De Melo-Martín and Intemann (2018, 4) likewise call consensus a “benchmark for scientific knowledge,” while emphasizing that dissent helps to test what grounds it: consensus views are more likely to be reliable when they “survive critical scrutiny” (2018, 21). Solomon (2017, 353–54) brings this concern to medicine, distinguishing informal agreement that emerges through research and clinical practice from formal statements produced by consensus panels. Expert agreement can be a useful, fallible guide for non-experts, but its significance depends in part on who participated, whether dissenters were heard, and whether peer pressure shaped the result.

Disagreement, dissent, and criticism are therefore essential to the scientific process. They can expose overlooked assumptions, test prevailing views, and sometimes overturn them. This is why C-SCOPE seeks to map the full distribution of expert judgments within a scientific community rather than simply determine whether consensus exists. However, it is

important to note that while C-SCOPE can map the distribution of views within a scientific community, it cannot establish whether that community has the conditions for Longino's transformative interrogation or whether a consensus has been critically achieved in the way that philosophers of science have highlighted. Nor can its survey instrument determine whether the disagreement or dissent it records is epistemically warranted. These are limits of what a survey can establish: no instrument can settle such normative questions on its own. As de Melo-Martín and Intemann (2018) emphasize, distinguishing warranted from unwarranted dissent is extremely difficult in practice.

Once we move beyond the idea of consensus as a vote on truth, the more substantive question emerges: "What epistemic information is conveyed by the distribution of expert judgment within a scientific community?" This requires us to consider not simply whether a majority endorses a claim, but how widespread and stable that endorsement is, and where disagreement persists. Patterns of agreement and dissent can help reveal where uncertainty remains and how disciplinary differences shape the interpretation of evidence. They also provide a starting point for further philosophical and

sociological inquiry into how those patterns arose and what further epistemic significance they have. The presence of dissent may indicate that prevailing views remain open to criticism and may be evidence that a community is functioning well. Moving from a narrow focus on “consensus” to a more expansive analysis of community-level opinion thus opens a more nuanced and empirically grounded framework for understanding the epistemic significance of collective expert judgment.

3.2 Surveying whole communities doesn’t tell us very much

Dispelling the standard slogans clears the ground for a deeper question: granting that collective opinions can be epistemically meaningful, *whose* opinions matter? Some critics argue that surveying entire scientific communities is unnecessary because what really matters are the views of a small number of senior specialists—the “super-experts” or recognized authorities in a field (Vickers 2023; Lane 2025). This intuition has an obvious appeal. Expertise is unevenly distributed, and it is natural to assume that senior

researchers, or those working directly on topic X, are the most reliable guides to its evidential status. On this view, if we want to understand the state of play regarding X, we should simply ask the handful of scientists best placed to know.[7]

Valuing the judgments of a small group of experts does not rule out *also* surveying the broader community. This is simply a different kind of information, and it might well be worth having 'very broad' information (at the large community level), and 'quite broad' information (at a smaller community level), in addition to 'narrow' information. Several considerations support our claim that both narrow and broad approaches are valuable:

[7] Interestingly, Miriam Solomon (2015), in her study of NIH consensus conferences, casts doubt on the epistemic value of consulting only a small group of experts. Beginning in the 1970s until the program phased out in 2013, the NIH convened consensus panels designed to review scientific evidence and produce authoritative statements intended to guide medical practice and public understanding. Yet Solomon argues that these panels rarely generate genuinely new knowledge; instead, they typically formalize and communicate patterns of expert judgment that already exist within the relevant scientific community. Their authority, therefore, derives less from the deliberations of the panel itself than from the broader distribution of opinion and evidence within the field that the panel attempts to represent.

(i) Community-wide surveys capture features of scientific opinion that narrower expert polls often cannot. A study based on only a few dozen scientists provides little sense of how views are distributed across the wider field. For example, Myers et al. (2021) report 98.7% agreement on anthropogenic climate change, but this figure is derived from 153 carefully selected specialists drawn from a global community numbering in the tens of thousands. While such studies can provide valuable evidence about expert opinion, their narrow sampling raises understandable questions about representativeness and selection effects. Broader surveys do not replace expert-focused studies; rather, they contextualize them by situating specialized judgments within the larger distribution of opinion across a scientific community and reducing the appearance of cherry-picking.

(ii) Narrow expert groups are more susceptible to groupthink. Close-knit communities of senior researchers often share background assumptions, review each other's work, and reinforce the same conceptual frameworks (cf. Solomon 2006; Smolin 2008). These sociological dynamics are an inevitable by-product of specialization. Expanding the sample reduces these effects. Only by adopting an extremely skeptical view—

treating the entire international community as a single coordinated echo-chamber—could one argue that broad surveys simply reproduce groupthink at scale.

(iii) History shows that small clusters of senior experts can be systematically mistaken. Continental drift is an instructive case. The most prominent geologists in North America dismissed drift with remarkable confidence, yet scientists in other regions saw the evidence differently; there are some shocking quotes documented (Frankel 2012), and surely there are far more undocumented. The senior experts were not only wrong but entrenched in their views. Researchers deeply embedded within a specialist domain may lose some degree of critical distance. It is important to control for different sources of error, so having a more diverse group is more effective.

(iv) Broad surveys allow for the expression of uncertainty and facilitate methodological refinement. It is worth bearing in mind that scientists often self-select when asked a survey question: if they feel that they do not have an informed opinion, they can choose to (a) not vote, or (b) vote neutral.

Many scientists do not like to guess or speculate.[8] In addition, C-SCOPE could easily conduct revealing experiments; for example, surveying scientists on issues far from their area of expertise to see how willing they are to vote on things they know little about. An interesting experiment would divide a scientific community into two groups, one faced with a traditional 5-point Likert scale including 'neutral', and another faced with a 4-point or 6-point Likert scale lacking 'neutral' (forcing a decision). Data from such experiments could then be used to inform how surveys are designed, and how data are interpreted.

In short, we concede that consulting top experts is indispensable, but we argue that it is not sufficient. Broader surveys offer epistemic advantages that narrow approaches cannot match: they mitigate local biases, reveal structural

[8] See for example Hansen (2007, 5), "Glaciologist Richard Alley, an IPCC lead author, said 'Lots of people were saying we [IPCC authors] should extrapolate into the future, but we dug our heels in at the IPCC and said that we don't know enough to give an answer'". This kind of cautious attitude is typical of scientists generally, we suggest. In a study by Morgan and Keith (1995), 16 climate scientists were interviewed on a range of topics; it is evident that, rather than guess, they sometimes chose not to answer, in one case (Ibid., 470) stating explicitly that in their view others should not guess either, because the answer is simply unknown.

patterns of agreement and dissent, and capture how views are distributed across an entire field. Taken together, narrow *and* broad data provide a more complete and reliable picture of scientific opinion than either could alone.

3.3 Measuring or publicizing consensus promotes conformity

Another common concern is that emphasizing scientific consensus may homogenize thought within scientific communities by encouraging conformity and discouraging dissent. But this worry presupposes that projects such as C-SCOPE are primarily concerned with reinforcing consensus views. That assumption is misleading. Recent C-SCOPE projects concerning the possible existence and detection of extraterrestrial life (Vickers et al. 2025; Vickers et al. 2026), for example, are not primarily concerned with establishing whether consensus exists. The rebranding from the ‘Institute for Ascertaining Scientific Consensus’ (IASC) to C-SCOPE reflects our deliberate shift away from consensus as the sole object of interest and toward the broader distribution of opinion within scientific communities.

Consider a case in which many scientists assume that there is widespread consensus that X, but a survey reveals that expert judgment is in fact more divided and heterogeneous than commonly believed. In such a situation, making the distribution of opinion visible may reduce conformity rather than increase it. Scientists who previously believed their doubts were isolated may become more willing to voice criticism or reconsider dominant assumptions.

The history of continental drift provides a useful illustration. Between 1930 and 1950, many North American scientists assumed there was overwhelming consensus that continents could not drift (Frankel 2012). In a famous statement, the paleontologist George Gaylord Simpson wrote that “[t]he verdict of palaeontologists is practically unanimous: almost all agree in opposing [drift]” (Simpson 1943, p. 2). Yet this perception was misleading. Had a large-scale, anonymous, international survey been conducted at the time, it likely would have revealed substantial support for drift elsewhere, especially in South Africa. Awareness of this broader distribution of expert judgment may have prompted some North American scientists to re-examine the evidence more

seriously, perhaps even accelerating acceptance of what later became the dominant view.

The continental drift case thus illustrates both the conformity worry and a possible response to it. An apparent consensus may discourage dissent, while revealing the wider distribution of views may help ensure that a heterodox position receives scrutiny before it is dismissed. When a genuine consensus exists, it does not follow that publicizing it will simply encourage obedient conformity. In science, there are often significant intellectual and professional rewards for successfully challenging a dominant view.[9] Once scientists are made aware that a mainstream view is widely held, some enterprising maverick may be *more* motivated to scrutinize it, precisely because overturning a consensus yields high epistemic and professional rewards. Being the person who

[9] In recent years, philosophers of science have renewed their interest in the different ways scientists contribute to the scientific enterprise. Drawing on earlier insights from Kuhn (1996) and Hull (1988), we can distinguish between *mavericks*—those who challenge conventions, break rules, and push research in novel directions—and *followers*, who work within established paradigms and contribute through the steady practice of normal science. Various philosophers have explored the incentives that drive scientists to become mavericks rather than followers (Weisberg and Muldoon 2009, Thoma 2015, Pöyhönen 2017, Heesen 2019).

successfully disconfirms, refines, or reconceptualizes a consensus position can be career-defining. Transparency about patterns of expert judgment may therefore generate not conformity, but strategic and productive dissent (Santana 2021).

Of course, it remains possible that some scientists would respond to survey data by “jumping on the bandwagon” of whichever view appears dominant. In her YouTube critique of Vickers et al. (2024), Hossenfelder (2024) raises precisely this concern. But bandwagon effects have long existed within science independently of attempts to measure expert judgment systematically. For example, Feyerabend (1996, 77fn.) reports on the fact that scientists in the 1950s uncritically accepted Von Neumann’s supposed proof that a ‘hidden variables’ interpretation of quantum mechanics was impossible. This idea spread around the community – people heard that Von Neumann had proved the result, and didn’t bother to scrutinize the proof. But such dynamics have long been part of scientific life. It is not at all clear that the existence of data on the distribution of opinion within scientific communities would make this tendency more powerful, rather

than less powerful—or that it would have any significant effect at all.[10]

4. Communication concerns

Even if data outputs are epistemically defensible, problems could still arise in how the findings are communicated and interpreted. These objections focus on how expert judgment figures might be presented to the public, whether the project inadvertently reproduces a ‘deficit model’ of science communication, and whether the data might be misused or misunderstood by various audiences.

[10] It has been suggested that making scientific agreement or disagreement more visible could distort research funding decisions. In practice, however, any such effects are likely to be limited. For example, funding agencies’ preference for hypothesis-driven work appears to arise primarily from administrative and evaluative constraints rather than from perceptions of consensus as such (Haufe 2013). As a result, even where community-level (dis)agreement is documented, funding behavior is unlikely to shift substantially. At most, such data might serve as a supplementary input for funders explicitly seeking to encourage epistemic diversity (Kitcher 1990), though whether this would occur remains an open empirical question.

4.1 The data produced could result in lowering public trust in science

C-SCOPE’s results may reveal significantly different levels of agreement from those expected on a given topic, including substantial disagreement among experts, potentially undermining public trust in science. C-SCOPE could also ‘tread on the toes’ of pre-existing consensus results and potentially undermine or contradict their results in problematic ways. For example, imagine C-SCOPE conducting a survey on whether climate change is anthropogenically caused. It is expected, given previous studies (Cook et al. 2013; Myers et al. 2021), that a high degree of consensus would be achieved. However, C-SCOPE's results may present lower degrees of expert agreement given different surveying methodologies, e.g. choosing to survey a very large, broad sample rather than smaller, curated groups of highly specialized experts. Even small disparities in data results could be exploited to undermine public trust (Lewandowsky et al. 2013; Maertens et al. 2020), especially as bad actors look to exploit any degree of dissent as evidence of uncertainty and a lack of confidence (Oreskes and Conway 2010). While it may be obvious to some that different surveying methodologies will naturally turn out

different consensus scores, this could be easily overwhelmed by malicious actors leveraging the unvetted nature and ease of access of social media.

A large proportion of the public perceive science as being about proving theories with certainty, or as a “search for truth” (Rabinovich and Morton 2012; Bertoldo et al. 2019, 788). In fact, most students hold a form of ‘naïve scientism’ where, as they see it, stable universally accepted truths are obtained via rigorous application of the scientific method (Tosoni and Ricotti 2024). Some scientists themselves attempt to maintain this kind of public-facing image of science, out of concerns for public distrust, by deliberately simplifying or obscuring scientific uncertainties (Zehr 2000, 95-97). However, this often means that when scientists discover new evidence which updates their current theories the public can perceive this as a lack of reliability and as contradicting results (Intemann 2023, 356). This, in turn, leads to public trust in science being undermined (John 2018, 82) as confidence in science’s capacity to access the ‘truth’ is lowered. Low degrees of agreement on scientific topics could make the public think that scientists don’t know what they are

talking about (Intemann 2023, 356-7), rather than recognizing that science is a process defined by discussion and debate.

We can restate this problem in terms of a passive and an active sense. The passive problem is that results showing expert disagreement could be interpreted as uncertainty, incompetence, or unreliability of scientists. The public loses confidence in science when faced with conflicting expertise among scientists; C-SCOPE, by producing results that reveal conflict within the scientific community, could therefore (allegedly) fuel this problem. The passive problem can be actively worsened by bad actors engaging in agnotology (Proctor and Schiebinger 2008) strategically amplifying and highlighting uncertainty to sow doubt about the nature and reliability of scientific knowledge. In the cases of both the tobacco and fossil-fuel industries, strong scientific consensuses were undermined by the mere scope for doubt and its exploitation (Oreskes and Conway 2010). Expert disagreement results from C-SCOPE could give easy fuel to merchants of doubt and exacerbate public distrust in science or at least worsen public support for public policies which are in the interests of political or financial elites.

The criticism hinges on the following premise: the public's lack of understanding of the nature of science results in distrust or loss of confidence in scientific results when they are presented with high degrees of uncertainty or internal conflict in science.

Determining whether communicating scientific uncertainty undermines trust is an empirical issue, and the evidence is mixed and context-sensitive. Importantly, different kinds of uncertainty appear to have different effects: communicating disagreement among scientists may be more likely to reduce credibility than communicating technical uncertainty such as error ranges or probabilities (Gustafson and Rice 2020).[11] Since C-SCOPE may often make precisely this kind of scientific disagreement salient, the possibility of a trust-reducing effect should therefore be taken seriously. We must consider how the project fares whether or not such effects occur.

To the extent that such trust-reducing effects arise from expectations that science should deliver certainty and unanimity, one appropriate longer-term response would be a shift in public education and science media reporting toward

[11] See Johnson and Slovic (1995) for evidence that uncertainty perceived as reflecting incompetence can decrease trust, and Pidgeon and Fischhoff (2013) for discussion of how appropriately contextualized uncertainty can instead foster trust.

an understanding of science as involving debate, discussion, and disagreement, as has been much discussed in recent literature (Tosoni and Ricotti 2024; Towler 2025).

One possible response to this risk would be to withhold or downplay evidence of scientific uncertainty or disagreement. That response, however, risks adopting a paternalistic conception of the public which has long been critiqued in STS (Wynne 1992; Jasanoff 2003c) assuming publics cannot grasp the social dimensions of science. STS scholars have long advocated for replacing the deficit model – which reduces public skepticism to ignorance – with a Nature of Science (NoS) framework (Lederman 2013; Towler 2025). NoS teaches that science is iterative, socially embedded, and thrives on debate. Yet scientists often obscure these dynamics, fearing public backlash (Zehr 2000), perpetuating a vicious cycle where hiding uncertainty reinforces the myth of infallibility, leaving publics unprepared for the reality of scientific practice. C-SCOPE's surveys could help break this cycle by exposing expert disagreement and normalizing the image of scientific contestation. It could offer a window into science's 'engine room' (Hilgartner 2000), demystifying how consensuses and scientific knowledge emerge from discourse. It could likewise

provide a form of ‘honest brokerage’ (Pielke 2007) which further antidotes distrust fueled by merchants of doubt through promoting institutional transparency conducive to trust-building (Pidgeon and Fischhoff 2013) and a public understanding of science founded on trust in science, due to understanding its human dimensions (Jasanoff 2003c; Pielke 2007). This capacity to help shift the norms of public understanding of science helps combat both the active and passive versions of this problem. The passive is taken to be a result of this public understanding of science which C-SCOPE can help change. The active version has bad actors deliberately exploiting this tendency of poor understanding and fueling doubt to further reinforce the distrust. Both have their capacities for public harm diminished by reaffirming NoS.

However, this remains a speculative defense and a possible component of a longer-term solution, without eliminating the potential for harms in the interim. Because the magnitude and persistence of any trust-reducing effects are likely to depend on context and presentation, C-SCOPE should treat them as a genuine communication risk rather than assume either that revealing disagreement will undermine trust or that transparency will necessarily strengthen it.

The issue can therefore be reframed as one of research ethics: given the possibility of social harm, is the production and dissemination of C-SCOPE data justified? This question is taken up further in Section 5.2. The statement-selection procedures outlined in Section 2 nevertheless provide a mechanism for identifying and avoiding particularly problematic cases, such as rerunning well-established consensus results, heavily value-laden questions (e.g. ‘safety’), and questions that are themselves ethically problematic to ask (e.g. ‘are there differences in intelligence between races?’). This approach to statement selection recognizes that patterns of scientific agreement and disagreement are especially relevant to public trust in areas of science that directly intersect with citizens’ lives. Many areas of research for which interesting and illuminating data about such (dis)agreement structures of the scientific community won’t have direct bearing on public-facing issues – i.e., science falling outside of public health and the environment. This isn’t to say C-SCOPE can or will never touch these more politically salient areas of science, but when this does occur, C-SCOPE, through its steering committee and its location within an academic research environment with access to

relevant expertise in sociology, science communication, and social analysis, will be well placed to act as an honest broker (Pielke 2007) actively managing this interface of knowledge, its public dissemination, uptake, and use.

4.2 This is a form of the much maligned ‘deficit model’

The deficit model holds that public resistance to science stems chiefly from a lack of knowledge; provide more facts, and acceptance follows. This picture has been widely criticized in STS and science communication (Wynne 1992; Jasanoff 2003c; Sturgis and Allum 2004; Irwin 2006; Wynne 2008) for oversimplifying how people form views – neglecting values, identity, trust, and lived experience – and for encouraging one-way, top-down ‘information dumping’ as opposed to two-way dialogue and engagement. C-SCOPE’s provision of data on community opinion might initially seem like just such ‘information dumping’ which, although not asserting scientific claims, may seem like an indirect attempt to influence the public’s beliefs in a top-down way, essentially another form of the deficit model.

The crucial distinction is that C-SCOPE's institutional role is descriptive rather than didactic. Its methods (preregistration, sampling disclosure, open release of anonymized data) foreground how the numbers are generated, inviting scrutiny rather than asserting authority. The results are also intended for multiple audiences. Policymakers, journalists, educators, and researchers may find them useful in different ways. Importantly, every set of findings is presented with caveats about scope and limits: the community sampled, the precise wording of questions, the response rates, and any translation issues. This contextualization is intended to prevent overclaim and to make clear that C-SCOPE outputs are inputs to deliberation, not conclusions in themselves. Although such information may legitimately influence public beliefs, C-SCOPE does not seek to promote particular conclusions independently of what its surveys reveal.

Why, then, might some observers associate C-SCOPE with the deficit model? The answer lies in the way consensus figures are sometimes used. In public debate, they can be wielded as rhetorical cudgels ("The science is settled, therefore you must…") regardless of how carefully they were produced.[12]

[12] Cf. Winsberg et al. (2020) for criticism of policymakers' use of 'following the science' rhetoric to justify COVID-19 restrictions.

That tendency reflects a communication problem, not an inherent flaw in the act of measuring scientific opinion. We also stress that C-SCOPE results are likely to reveal a solid scientific consensus only exceptionally; surveys on the vast majority of scientific topics will reveal some not-insignificant measure of community disagreement.

To guard against deficit-style uses, C-SCOPE should commit to certain design principles. Results must report the full distribution of responses, including neutrals, rather than offering a single ‘consensus score’ in isolation. The scope of each survey should be clearly qualified: which community was sampled, what counts as ‘relevance’, what expertise boundaries are assumed, and so on. Reports should distinguish between scientific findings and value-laden policy choices, so that levels of agreement are not confused with prescriptions for action. Public communication should frame results as starting points for dialogue rather than as endpoints. And where multiple wordings or scales are tested, results should be published side by side, to show the sensitivity of findings to survey design. These kinds of design principles also help safeguard C-SCOPE against the criticisms

raised in Section 4.1. Much of this is already evident in Vickers et al. (2024) and Vickers et al. (2025).

Even with these safeguards, some actors will still present C-SCOPE outputs as though ‘more numbers’ alone should resolve public debate. The appropriate response, however, is not to withhold information but to ensure transparent communication and plural engagement.

5. Political concerns

Some critics have raised concerns about the broader political implications of C-SCOPE’s work, particularly regarding how data on distributions of expert judgment might shape public decision-making, policymaking, and the relationship between scientific expertise and democratic authority. These concerns often take the form of worries about technocracy or elitism—specifically, that a body dedicated to representing scientific opinions could grant disproportionate influence to experts in public life. This section addresses such objections by examining direct comparisons that have been drawn between C-SCOPE and Collins and Evans’ (2017) proposed institute of The Owls. Taking these comparisons as a starting point, we

clarify the key ways in which C-SCOPE differs from The Owls and show how these differences allow us to confront—and ultimately avoid—the political criticisms of technocracy and elitism that have been leveled at Collins and Evans, and which might otherwise be thought to be inherited by a similar body such as C-SCOPE.

5.1 C-SCOPE for the birds?

5.1.1 Collins, Evans and the Owls

A response repeatedly encountered is that C-SCOPE has a strong resemblance to the 'Owls' concept of Collins and Evans (2017).[13] The alleged consequence is that C-SCOPE inherits the same problems, most notably that it will also tend toward an elitist conception of science. We show here that this resemblance breaks down in certain important respects. Questions that still arise concerning whether the project is somehow 'technocratic' or 'elitist' will be considered in section 5.2.

To see the resemblance, take the following quote from C&E (2017, 86):

[13] From here we refer to work jointly authored by Collins and Evans as C&E.

“The job of The Owls would be to look at the current state of expert knowledge and pass their conclusions on to the politicians, for them to use or over-rule. The Owls, remember, would have the relatively easy task of reporting the consensus, not the hard, perhaps impossible, task of knowing the truth.”

Given our interest in the distribution of expert judgment across scientific communities and our awareness that policymakers might be interested in this, drawing a conclusion that our project is similar is not without merit. The Owls would be a group of experts who provide the government with advice on scientific issues of relevance to policy. It would be composed of both scientists and social scientists, functioning together to examine, evaluate and determine the current state of consensus on technical scientific topics. From here the Owls would be able to advise the government on the substance and degree of consensus on any given technical issue (Collins and Evans 2017, 91).[14] The degree of consensus

[14] See footnote 7 on NIH panels. It's worth noting the parallels between the Owls and the NIH panels. Both are comprised of small groups of experts whose role is to convey levels of scientific consensus on given topics to policymakers. As noted already, for more on these NIH consensus panels see Solomon (2015).

would be graded on a scale from A to E with near complete consensus within the relevant community graded at an A while little to no consensus is graded at an E (2017, 90).

Importantly, C&E envision the Owls to have a position within the legal and constitutional structure of the state, existing as a body with the function of representing scientific consensus to political decision makers. As C&E (2017, 91) put it, "*...they would be a statutory committee.*" In the UK context, this is a committee created by an Act of Parliament or other secondary legislation which gives the committee a legal mandate to perform a particular function. The key takeaway from this is that the Owls would not be a voluntary body or, say, an NGO. It would be a body legally mandated to fulfill the function that C&E (2017) envision. Their advice to the politicians would not be binding, politicians could overrule them, but they would have a legally or constitutionally ordained presence within the state, whose reports and conclusions on consensuses would always be a matter of public record (2017, 91); open for examination relative to the decisions that were then actually made by the politicians.

C&E take most scientists to be ‘eagles’ (Collins and Evans 2017, 78; Collins 2022, 879), ‘powerful predators’ narrowly focused on the pursuit of truth in their given discipline. These scientists are not philosophically or sociologically interested in reflective understandings of what science is, rather they are focused on the pursuit of evidence and truth. In contrast to the eagles are ‘owls’. This is a minority of scientists capable of looking at the big picture, able to turn their gaze backward to analyze their own discipline akin to the way philosophers and sociologists of science do. It is this latter type of scientist who would be included.

In addition to the (natural) scientists, there would be a complementary set of ‘owls’ from the social studies of science. These would be individuals who started out reflecting on the sociology of science earlier in their careers but then gained some technical understanding of scientific fields which they worked on (Collins 2022, 880). They would have what C&E term ‘interactional expertise’ within the scientific fields they have studied. Someone with interactional expertise can understand and speak the language of a technical community fluently – through immersion and social interaction – without possessing full contributory expertise, i.e., the ability to carry

out the technical work itself (Collins and Evans 2002, 243-4; Collins et al. 2016). C&E (2017, 91-92) justify the position of social scientists and sociologists within the Owls as bringing a metascience perspective to the body. This includes, for example, understanding the way scientific controversies unfold and understanding and enhancing the visibility of experience-based experts whose role might be overlooked by scientists acting alone.

5.1.2 Divergence from The Owls

Given the mutual interest in phenomena like scientific community agreement, and our intention to try to map the distribution of expert scientific judgment on topics, there is a putative resemblance between the Owls and C-SCOPE. However, there are at least two important points of divergence:

(i) *Descriptive vs Prescriptive:* The Owls are more prescriptive in terms of their relation to policy and

public decision making. Their graded consensuses are intended to weigh in on policy decision-making and to act as a source of informational anchoring in public deliberation on technical scientific issues that are related to policy decision making. Politicians can be retrospectively held to account relative to the advice they received from the Owls. Rejecting advice stemming from a grade A consensus, just as utilizing a grade E consensus as the justification of a decision, requires a high burden of explanation for those decisions on the politicians. C-SCOPE, however, is more on the descriptive end of the descriptive-prescriptive spectrum. It is aimed at providing snapshots of scientific community opinion. While this may be a useful source of information that can be utilized as an input into decision making, it is not as directly prescriptive in relation to policy decisions. This shows another difference between C-SCOPE and The Owls vis-a-vis their relationships with the political sphere. In this regard, C-SCOPE is not as technocratic as the Owls because it sits outside of the frame of being a statutory body. C-SCOPE is more focused on

the descriptive exercise of giving a distribution of expert judgment across relevant scientific communities. It is not claimed here that the 'mere informational' snapshot of scientific community opinion data is value free and bereft of potential political problems – these will be taken up in 5.2 below.

(ii) *Measuring the distribution of scientific judgment:* C-SCOPE has a clear method for measuring agreement, disagreement and opinion across scientific communities, one which (as sketched in Section 2) has already been usefully operationalized in three studies. C&E do not in fact give an account of how The Owls would determine the degree of consensus on a given topic. They give an A-E grading scale, but what they leave out is exactly how the Owls will determine those grades. As it stands this seems to be left to the collective judgment of those that compose the Owls, presumably leveraging their networks and disciplinary knowledge of the relevant fields. This further opens the Owls to accusations of technocracy and elitism, in a way that C-SCOPE does

not. The C-SCOPE pilot studies to-date already explore various methodological nuances here, some of which seem incompatible with the C&E approach. For example, C-SCOPE is more fleet of foot, with the ability to zoom in or out when necessary. It can rapidly put out surveys to ascertain information on the state of scientific opinion on a given statement or set of statements and provide data that allows determination of (i) specialist scientific expert judgment, (ii) proximal scientific expert judgment, and (iii) non-proximal expert opinion (i.e. university-based scientists in general). For instance, C-SCOPE could survey statements about influenza viruses within the specialist community of virologists who focus on influenza. At the same time, a much broader community of virologists could also be surveyed on the same statements. In short, C-SCOPE can cast its net wider than the Owls and not be reliant on a relatively small body of individuals drawn from an elite of the scientific and sociological community.

5.2 The power held by such a standing body would be technocratic and elitist

5.2.1 The Political Alignment Objection

A growing body of empirical work shows that communicating the existence of a strong scientific consensus—consensus messaging—can be effective in persuading the general public to adopt beliefs aligned with expert judgment (van der Linden 2021; Bartoš et al. 2022). Some studies, however, indicate that consensus messaging can produce backfire effects among subpopulations with particular political background beliefs, leading recipients to become more entrenched in contra-consensus views. Importantly, these studies also report a positive overall effect of consensus messaging on belief formation across the broader population (Deryugina and Shurchkov 2016, 12; Suldovsky and Akin 2023, 113). Other work finds that consensus messaging can reduce political polarization and opinion divergence rather than exacerbate it (van der Linden et al. 2018, 2). Consistent with these findings, meta-analytic evidence indicates a robust overall effect of consensus messaging as a persuasive intervention (Rode et al. 2025).

If these studies are to be believed, the data produced by C-SCOPE is likely to possess significant persuasive power and, therefore, influence the public's beliefs about scientific matters. Precisely because of this persuasive capacity, C-SCOPE can be understood as exercising a form of epistemic power with political consequences. Since public beliefs about scientific matters often function as inputs to democratic deliberation and policy justification, institutions capable of shaping those beliefs occupy a structurally significant position within democratic decision-making (cf. Estlund 2008; Kitcher 2011). It is, therefore, easy to imagine scenarios in which C-SCOPE becomes closely aligned with government action. Indeed, there has been an explicit ambition to assist policymakers in their decision-making (Vickers et al. 2024, Conclusion). On this model, C-SCOPE's role is not to prescribe particular policy outcomes but rather to supply policymakers with descriptive data to inform science-related decisions, such as public health interventions during a pandemic or strategies for climate change mitigation.

This combination of epistemic influence and policy relevance raises a familiar legitimacy worry. Complete political neutrality is unrealistic: individuals and institutions inevitably operate

against a background of values and assumptions. The problem, however, is not value-ladenness as such. Rather, even if no such alignment exists in practice, the perception that C-SCOPE is politically aligned may be sufficient to undermine its credibility. Suspicions of political bias can erode trust independently of actual misconduct. This is a paradigm case of external capture: even in the absence of formal control, governments or funders may shape what questions are asked, when evidence is produced, and how results are framed, generating a de facto alignment that compromises perceived neutrality.

The simplest, or ‘naïve’ version of this objection holds that C-SCOPE might deliberately manipulate distribution of expert judgment data to advance preferred agendas—either by fabricating the appearance of scientific consensus where none exists or by obscuring genuine consensus where it does. For example, critics might worry that, given the left-leaning political orientation common among academics, such manipulation might be used to exaggerate the appearance of expert agreement on policy-salient issues like climate change to bolster ambitious regulatory agendas. Such interference could, in principle, take several forms: (i) designing survey

instruments in ways that prime or nudge respondents toward particular answers;[15] (ii) restricting the surveyed population to scientists whose views fall within a predetermined domain;[16] or (iii) misrepresenting results, up to and including outright data fabrication.

These concerns, however, are best understood as limiting cases rather than the central legitimacy threat. Conceptually,

[15] There are legitimate methodological concerns about statement biases that could skew results (cf. Isoaho et al. 2019; Congiu and Moscati 2022, 204-207) especially given differences in interpretation in individuals' translations (cf. Pérez and Tavits 2017; Velez et al. 2024, 1222-24). These are, however, recognized openly (Vickers et al. 2024, 28-30) and will help inform C-SCOPE to minimize biasing effects in future. For the purposes of this criticism, we are considering a deliberate manipulation of statements to produce a desired consensus result.

[16] There are legitimate methodological concerns about sampling biases that could skew results, including disparities in academic funding (Ginther et al. 2018; Erosheva et al. 2020), the systematic underrepresentation of some demographic groups within STEM (Fry et al. 2021), and inequalities in academic career progression (Hofstra et al. 2020, 9287–9288). Such structural biases cannot straightforwardly be eliminated by survey design, and these broader methodological issues are beyond our present scope. For the purposes of this criticism, we are considering the deliberate manipulation of the domain of scientists surveyed in order to produce a desired consensus result.

and as noted in Section 2, C-SCOPE's primary aim is not public persuasion, but rather the systematic collection, analysis, and dissemination of information about the beliefs of scientific communities, and its methods are designed to be open to external scrutiny. Many of the manipulations envisaged here—particularly those involving survey design or misrepresentation—would be detectable by informed external observers under conditions of methodological transparency. The level of opacity required to conceal such practices would be difficult to sustain under the methodologies outlined in Section 2, and would itself undermine trust and invite suspicion among both scientists and informed external observers.

Importantly, even if the procedural safeguards discussed in Section 2 are sufficient to rule out bad-faith manipulation, they do not exhaust the legitimacy concerns raised by C-SCOPE's epistemic role in democratic societies.[17] The deeper

[17] This concern can be understood in Rawlsian terms as a question of public justification: whether the epistemic influence exercised by institutions like C-SCOPE can be justified to citizens as free and equal participants in democratic decision-making, rather than merely defended on epistemic grounds alone. On public justification and the legitimacy of political power, see Rawls (1996; 2001) and Gutmann and Thompson (1996; 2004).

worry is not that C-SCOPE will act improperly, but that reasonable methodological choices (e.g. how questions are framed, which populations are sampled, or how results are presented), made in good faith and under transparent conditions, may nonetheless shape public perceptions in ways that are not justifiable to all those subject to their influence, since reasonable people who reject the substantive findings may reasonably contest the legitimacy of those methodological choices.[18]

Even so, it is worth clarifying the appropriate standard against which such concerns should be evaluated. In particular, it is important to distinguish worries about potential misconduct from questions of political legitimacy. Any institution that occupies a position of epistemic or political influence is vulnerable, in principle, to misuse or abuse. That fact alone cannot serve as a decisive objection, unless one is prepared to adopt a level of skepticism that would undermine the

[18] We do not address here the further question of what institutional arrangements would be required to render such epistemic influence democratically legitimate. Addressing that question would require engaging political theories of epistemic authority, public justification, and the role of expertise in democratic societies in a way that goes beyond what is required to respond to the present objection.

legitimacy of all institutions that exercise power. Extreme skepticism of this kind is not only impracticable but normatively uninformative: it offers no principled basis for distinguishing acceptable from unacceptable forms of institutional authority. The central question, therefore, is not whether C-SCOPE's epistemic power could be abused—since this is true of any such institution—but whether that power is legitimate under conditions of transparency, accountability, and its proper role within democratic deliberation. Framed in this way, the relevant task is to assess the legitimacy of C-SCOPE's role, rather than to demand guarantees against all conceivable misuse. Fully undertaking that task, however, would require a much more extensive treatment than can be provided here and therefore lies beyond the scope of the present paper.

5.2.2 The Technocratic Elitist Objection

The spirit of the naïve objection nonetheless gains traction when considered against the backdrop of the "following the science" rhetoric that accompanied policy justification during the COVID-19 pandemic (Cairney 2021; Evans 2022). In that context, the views of a relatively small and select group of

experts were often treated as standing in for ‘Science’ itself and were invoked to legitimate political decisions (Winsberg et al. 2020). Given that the public generally accords science a privileged epistemic status (Bertoldo et al. 2019, 788–91), and given the persuasive force of communicated consensus, it may be argued that C-SCOPE’s results could similarly be mobilized to justify technocratic forms of decision-making. On this view, the objection to C-SCOPE is that it risks functioning as a problematic tool of technocracy.[19]

This concern rests on assumptions that do not follow from C-SCOPE’s institutional role. First, it presupposes that C-SCOPE will be tightly integrated into policy advising in a way that systematically serves policymakers’ preferred outcomes. But as already discussed in 5.1, C-SCOPE would not be so

[19] There are broadly two conceptions of technocracy in public decision making: (i) extreme technocracy, where science replaces politics as decision-maker, organizing society according to whatever ideals the evidence supports; (ii) milder technocracy, where an institutionalized division of epistemic labor means policymakers work within constraints set by technical experts, choosing between options they define (Krimsky, 1984). (i) is a 'bogeyman' conception unlikely to be seen in current democracies. (ii) is a more legitimate concern, stemming from governments' increasing reliance on technical expert advisors (cf. Wynne, 1992; 2008), and is the form we have in mind here.

integrated. Second, the objection assumes that C-SCOPE's surveys will typically yield strong, policy-relevant collective expert judgments. Yet it is at least as plausible that many surveys will reveal substantial disagreement within the scientific community (as already seen in the pilot studies and discussed in Section 3.3). High levels of collective expert judgment sufficient to ground policy justification are hardly guaranteed. Indeed, a persistent challenge for science-informed policymaking is that political decisions are often required precisely when scientific expert agreement has not yet emerged (Collins and Evans 2002, 239–40). In such cases, the proper role of political representatives is not to defer judgment until uncertainty is resolved, but to assume responsibility for decision-making under conditions of uncertainty. Systematically documenting the extent and structure of scientific disagreement—that is, the absence of expert agreement—can therefore inform political deliberation by making uncertainty publicly transparent, while leaving responsibility—and the associated risk—squarely with elected decision-makers.

There is an important limitation here. C-SCOPE will not be equally well suited to every scientific claim bearing upon public decision-making. Some highly policy-salient questions

are interdisciplinary or inseparably entangled with normative considerations concerning acceptable risk, costs, or competing social priorities. In particular, C-SCOPE is not intended as an instrument for resolving fine-grained, all-things-considered policy questions.

This does not imply either that C-SCOPE is confined to foundational scientific claims or that more upstream scientific questions lack policy relevance. Many apparently downstream policy questions combine a number of empirical questions concerning likely effects, risks, and consequences with a further judgment about how those considerations should be weighed. C-SCOPE can map scientific judgment concerning the former without purporting to resolve the latter. Policymakers may therefore have good reason to know where a relevant scientific community stands on empirical matters even when those matters substantially underdetermine policy. A longitudinal survey of a scientific community's judgment concerning a question such as whether global warming is likely to exceed 2°C by 2100, for example, could provide useful information about the development of scientific opinion without implying any particular judgment about mitigation, adaptation, or other policy choices.[20] Where an overall policy

[20] Precise wording might be difficult to agree, of course; use of the word ‘likely’ could be controversial, since it could be interpreted differently by different participants.

question combines empirical and normative elements, C-SCOPE's appropriate role is therefore to map expert judgment concerning the empirically tractable components of that question, leaving the further question of how those considerations should be weighed to the political process.

This division of labor is captured by Collins and Evans' distinction between the technical and political phases of decision-making (Collins and Evans 2002; 2007; 2017). On this account, institutions like C-SCOPE contribute to the technical phase by supplying information about the state of scientific opinion. Decisions about whether and how to act on that information properly belong to the political phase, which is occupied by citizens and their representatives.

Another branch of this criticism may arise here, given that C-SCOPE provides its findings to democratic deliberation during the 'technical phase,' thereby perpetuating an elitist epistemic division of authority. In virtue of being epistemically deferred to, scientists' views are granted disproportionately greater power to inform policy decisions (Wynne 1982; Jasanoff 1997; Irwin 2002) and to determine the meanings, frameworks, and salient questions relevant within 'technically-involved' democratic decision making. This criticism of C-SCOPE follows

the historical charges of elitism against C&E's technical-political phase distinction (Jasanoff 2003a; Rip 2003; Wynne 2003) and STS charges against epistemic deference to traditionally conceived and institutionally ingrained sources of expertise (Wynne 1992; Epstein 1995; Irwin 2002; Ottinger 2013). The thought here is that by systematically representing the views of credentialed scientists, C-SCOPE privileges a particular epistemic community and may thereby reinforce institutional scientistic framings of public problems—such as how key terms like 'risk,' 'safety,' or 'evidence' are defined—at the expense of wider public perspectives (Wynne 1992; Jasanoff 2003a, 2003b; Wynne 2003, 2008). However, this concern does not arise from the mere use of scientific expertise, which is an unavoidable feature of the division of epistemic labor in complex societies (Hardwig 1985; Collins and Pinch 2008), but from the potential influence such framings may exert over how problems are structured and interpreted during democratic deliberations.

These STS concerns about epistemic elitism are not unfounded, as many of these critiques emerged in a context in which scientific authority plausibly appeared to marginalize lay perspectives (Wynne 1992; Epstein 1995; Irwin 2002;

Wynne 2003). This epistemic elitism was not only a source of epistemic injustice (Fricker 2007) but also epistemically exclusionary at the expense of mutually beneficial expansions of their relevant epistemic domains. However, in our modern epistemic and political contexts it is not clear that STS's more skeptical approach to expertise is sustainable. In contemporary public discourse, scientists are often not the most powerful actors. Rather, they are frequently overshadowed by political, corporate, and mainstream and social media actors with substantially greater influence. Although many scientists are socially and perhaps economically privileged, privilege alone does not determine epistemic authority.

A more salient feature of the current sociopolitical environment is the erosion of scientific authority in public life, rather than its overextension. In this context, treating all expertise as equally suspect risks exacerbating the very epistemic vulnerabilities that STS scholars often seek to diagnose. Persistent critiques of elitism often reflect a democratic preference for broad public participation in framing policy problems, even where those problems involve highly technical matters. We take the opposing view: in

technical domains, democratic deliberation is weakened, not strengthened, when the epistemic authority of relevant expertise is unduly discounted. This is ultimately a disagreement in democratic theory.[21] It cannot be resolved here, but it does not by itself undermine C-SCOPE's legitimacy as a provider of technical input within democratic deliberation. In this role, C-SCOPE would not displace democratic judgment but function as one epistemic input among many. When scientific opinion remains divided, political representatives must still exercise judgment under conditions of uncertainty rather than defer decision-making to technical expertise.

Importantly, the technocratic and elitist dangers lie not in the availability of expert information as such, but in its exclusive or uncontestable use. In this respect, C-SCOPE's outputs may mitigate rather than exacerbate technocratic risks. By making the degree of agreement – and disagreement – within a scientific community visible, C-SCOPE can expand the epistemic field beyond small, closed advisory groups and

[21] See Durant (2011) for how Collins and Evans' prioritization of expert authority on technical matters reflects a Rawlsian conception of democratic deliberation, while prioritizing public participation – as Jasanoff and Wynne do – reflects a more Habermasian and participation-oriented model of democratic deliberation.

render scientific claims more transparent and contestable.[22] In doing so, it can reduce the scope for the strategic invocation of “science” as a rhetorical cudgel wielded by either technocrats or demagogues (Collins et al. 2020, ch. 6). Access to structured information about scientific disagreement can thus function as a partial check on both technocratic overreach and populist epistemic monopolization (Collins et al. 2023).

Overall, this objection overstates the inherent political power of C-SCOPE and underestimates the mediating role of democratic institutions. The results of a C-SCOPE survey would contribute to, rather than determine, policy formation. They are, after all, one input among many. The epistemic authority of C-SCOPE’s project derives from its transparency and representativeness. Properly functioning democratic deliberation requires reliable technical input to inform the

[22] This can be seen in how the consensus presented by scientific advisory bodies during the pandemic was non-representative of those actually working on the relevant models. A curated selection of modelers' work, chosen by a small group of government advisors, was used to politically justify decisions made on the basis of 'following the science' – at a point when that work was outdated and no longer supported by its original authors (Rhodes and Lancaster, 2022). A body such as C-SCOPE would have the capacity to challenge this kind of technocratic rule.

political phase of decision-making, but the authority to act remains political, not scientific. As a properly constrained, non-coercive epistemic influence, aimed at clarifying the state of scientific agreement and disagreement, C-SCOPE may not merely be permissible but positively valuable in enabling citizens and policymakers to deliberate about science-related matters. In this sense, C-SCOPE is not a technocratic instrument but a democratizing one. Far from empowering experts to rule, C-SCOPE could enable citizens and policymakers alike to see where science speaks with one voice, and where it does not. Ultimately, the authority to act on these inputs lies with society's political representatives.

6. Taking stock

This paper takes seriously the notion of having a dedicated body for mapping the distribution of scientific expert judgment, namely, as embodied in the methodological and institutional structure of C-SCOPE. In doing so we have articulated a range of objections from various disciplines. While many of these concerns extend beyond the scope of the present paper, they do not undermine the broader project.

Rather, in engaging our critics, we have clarified the methodological, conceptual, and normative challenges that any serious attempt to study scientific judgment must confront. Taken together, these considerations provide a strong case for the existence of a standing body like C-SCOPE.

In our view, underlying the skepticism toward scientific consensus and its measurement is a view that social processes have little or no bearing on scientific claims. Popper's vision of science, centered on falsification and the lone theorist challenging orthodoxy, has encouraged a kind of romantic individualism: the belief that genuine progress arises from dissent rather than agreement. For some, it is therefore detrimental to science to admit any social negotiation. But this is a fundamental misunderstanding of how scientific inquiry proceeds. Patterns of scientific agreement and disagreement reflect collective expert judgment, rather than mere social consensus.

For others, their concerns lie in dangers around the reification of scientific expertise, risking the overextension of its legitimate use in democratic societies. Much of what we attempt to provide in response to these concerns takes the form of showing that C-SCOPE will not result in an

overextension of scientific expertise beyond its bounds. Our responses ameliorate these concerns by demonstrating how C-SCOPE's responsible institutional design tempers the potential for democratic overreach and instead enables greater democratization of scientific expertise beyond closed technocratic elites.

Our primary philosophical motivation for C-SCOPE is the need for systematic information about the distribution of community-level expert judgment. The evidence alone rarely determines exactly which inferences can be drawn. At any given time, almost any body of evidence can support more than one claim, and perhaps even multiple claims. So too, at a given time, a body of evidence may warrant belief in the opinion of some scientists, but agnosticism in the opinion of other (equally qualified) scientists. This is why expert judgment is required for science to proceed. Scientific experts can determine which theories are more likely than others given the same body of evidence. But experts can disagree, especially when the evidence is highly complex and uncertain (as it usually is). Thus, we require not only expert judgment to guide our interpretation of evidence but collective expert judgment. It is not the opinion of one scientist – even a very

good one – that matters, but the opinion of a community of scientists who are in dialogue with each other. This does not mean that collective expert judgment is infallible; rather, it is often our best bet in the face of uncertainty.

Despite centuries of organized science, we still lack any systematic record of how scientific communities as a whole position themselves relative to their fields of inquiry. It is remarkable that humanity can map the stars and the genome, yet not its own collective scientific judgment. The C-SCOPE project begins to change that.

References

Allum, N., A. Reid, M. Bidoglia, G. Gaskell, N. Aubert-Bonn, I. Buljan, S. Fuglsang, et al. 2023. “Researchers on Research Integrity: A Survey of European and American Researchers.” *F1000Research* 12:187.

Bartoš, V., M. Bauer, J. Cahlíková, and J. Chytilová. 2022. “Communicating Doctors’ Consensus Persistently Increases COVID-19 Vaccinations.” *Nature* 606 (7914):542–49.

Bergin, A. E., and H. H. Strupp. 1972. *Changing Frontiers in the Science of Psychotherapy*. Chicago: Aldine-Atherton.

Bertoldo, R., C. Mays, G. Böhm, W. Poortinga, M. Poumadère, E. Tvinnereim, A. Arnold, K. Steentjes, and N. Pidgeon. 2019. “Scientific Truth or Debate: On the Link between Perceived Scientific Consensus and Belief in Anthropogenic Climate Change.” *Public Understanding of Science* 28 (7):778–96.

Brown, R. F., A. S. John, Y. Hu, and G. Sandhu. 2024. “Differential Electronic Survey Response: Does Survey Fatigue Affect Everyone Equally?” *Journal of Surgical Research* 294:191–97.

Cairney, P. 2021. “The UK Government’s COVID-19 Policy: What Does ‘Guided by the Science’ Mean in Practice?” *Frontiers in Political Science* 3:624068.

Collins, H. 2022. “The Owls: Some Difficulties in Judging Scientific Consensus.” *Villanova Law Review* 67 (5):877–82.

Collins, H., and R. Evans. 2002. “The Third Wave of Science Studies: Studies of Expertise and Experience.” *Social Studies of Science* 32 (2):235–96.

Collins, H., and R. Evans. 2007. *Rethinking Expertise*. Chicago: University of Chicago Press.

Collins, H., and R. Evans. 2017. *Why Democracies Need Science*. Cambridge: Polity Press.

Collins, H., R. Evans, D. Durant, and M. Weinel. 2020. *Experts and the Will of the People: Society, Populism and Science*. Cham, Switzerland: Palgrave Macmillan.

Collins, H., R. Evans, D. Durant, and M. Weinel. 2023. “The Third Wave and Populism: Scientific Expertise as a Check and Balance.” Pp. 52–75 in *The Oxford Handbook of Expertise and Democratic Politics*. Edited by G. Eyal and T. Medvetz. Oxford: Oxford University Press.

Collins, H., R. Evans, and M. Weinel. 2016. “Expertise Revisited, Part II: Contributory Expertise.” *Studies in History and Philosophy of Science Part A* 56:103–10.

Collins, H., and T. Pinch. 2008. *Dr. Golem: How to Think about Medicine*. Chicago: University of Chicago Press.

Congiu, L., and I. Moscati. 2022. “A Review of Nudges: Definitions, Justifications, Effectiveness.” *Journal of Economic Surveys* 36 (1):188–213.

Cook, J., D. Nuccitelli, S. A. Green, M. Richardson, B. Winkler, R. Painting, R. Way, P. Jacobs, and A. Skuce. 2013. “Quantifying the Consensus on Anthropogenic Global Warming in the Scientific Literature.” *Environmental Research Letters* 8 (2):024024.

Cook, J., N. Oreskes, P. T. Doran, W. R. L. Anderegg, B. Verheggen, E. W. Maibach, J. S. Carlton, et al. 2016. “Consensus on Consensus: A Synthesis of Consensus Estimates on Human-Caused Global Warming.” *Environmental Research Letters* 11 (4):048002.

Curnow, T. 2012. *A Practical Guide to Philosophy for Everyday Life: See the Bigger Picture*. London: Icon Books.

de Koning, R., A. Egiz, J. Kotecha, A. C. Ciuculete, S. Z. Y. Ooi, N. D. A. Bankole, J. Erhabor, et al. 2021. “Survey Fatigue during the COVID-19 Pandemic: An Analysis of Neurosurgery Survey Response Rates.” *Frontiers in Surgery* 8:690680.

Dellsén, F. 2025. “Consensus versus Unanimity: Which Carries More Weight?” *British Journal for the Philosophy of Science* 76 (2):373–95.

de Melo-Martín, I., and K. Intemann. 2018. *The Fight against Doubt: How to Bridge the Gap between Scientists and the Public*. Oxford: Oxford University Press.

Deryugina, T., and O. Shurchkov. 2016. “The Effect of Information Provision on Public Consensus about Climate Change.” *PLOS ONE* 11 (4):e0151469.

Durant, D. 2011. “Models of Democracy in Social Studies of Science.” *Social Studies of Science* 41 (5):691–714.

Elsdon-Baker, F. 2009. *The Selfish Genius: How Richard Dawkins Rewrote Darwin’s Legacy*. London: Icon Books.

Epstein, S. 1995. “The Construction of Lay Expertise: AIDS Activism and the Forging of Credibility in the Reform of Clinical Trials.” *Science, Technology, & Human Values* 20 (4):408–37.

Erosheva, E. A., S. Grant, M. C. Chen, M. D. Lindner, R. K. Nakamura, and C. J. Lee. 2020. “NIH Peer Review: Criterion Scores Completely Account for Racial Disparities in Overall Impact Scores.” *Science Advances* 6 (23):eaaz4868.

Estlund, D. 2008. *Democratic Authority*. Princeton, NJ: Princeton University Press.

Evans, R. 2022. “SAGE Advice and Political Decision-Making: ‘Following the Science’ in Times of Epistemic Uncertainty.” *Social Studies of Science* 52 (1):53–78.

Feyerabend, P. 1996. *Killing Time: The Autobiography of Paul Feyerabend*. Chicago: University of Chicago Press.

Frankel, H. R. 2012. *The Continental Drift Controversy*. Cambridge: Cambridge University Press.

Fricker, M. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.

Fry, R., B. Kennedy, and C. Funk. 2021. *STEM Jobs See Uneven Progress in Increasing Gender, Racial, and Ethnic Diversity*. Washington, DC: Pew Research Center.

Ginther, D. K., J. Basner, U. Jensen, J. Schnell, R. Kington, and W. T. Schaffer. 2018. “Publications as Predictors of Racial and Ethnic Differences in NIH Research Awards.” *PLOS ONE* 13 (11):e0205929.

Greenhalgh, T., M. B. Schmid, T. Czypionka, D. Bassler, and L. Gruer. 2020. “Face Masks for the Public during the COVID-19 Crisis.” *BMJ* 369:m1435.

Gustafson, A., and R. E. Rice. 2020. “A Review of the Effects of Uncertainty in Public Science Communication.” *Public Understanding of Science* 29 (6):614–33.

Gutmann, A., and D. Thompson. 1996. *Democracy and Disagreement*. Cambridge, MA: Harvard University Press.

Gutmann, A., and D. Thompson. 2004. *Why Deliberative Democracy?* Princeton, NJ: Princeton University Press.

Hansen, J. E. 2007. “Scientific Reticence and Sea Level Rise.” *Environmental Research Letters* 2:024002.

Hardwig, J. 1985. “Epistemic Dependence.” *Journal of Philosophy* 82 (7):335–49.

Haufe, C. 2013. “Why Do Funding Agencies Favor Hypothesis Testing?” *Studies in History and Philosophy of Science* 44:363–74.

Heesen, R. 2019. “The Credit Incentive to Be a Maverick.” *Studies in History and Philosophy of Science Part A* 76:5–12.

Hilgartner, S. 2000. *Science on Stage: Expert Advice as Public Drama*. Stanford, CA: Stanford University Press.

Hofstra, B., V. V. Kulkarni, S. Munoz-Najar Galvez, B. He, D. Jurafsky, and D. A. McFarland. 2020. “The Diversity–Innovation Paradox in Science.” *Proceedings of the National Academy of Sciences* 117 (17):9284–91.

Hossenfelder, S. 2024. “Sociologists Want to Establish Scientific Consensus.” YouTube video, December 22. https://www.youtube.com/watch?v=zucXnn64qtk.

Hull, D. L. 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.

Intemann, K. 2023. “Science Communication and Public Trust in Science.” *Interdisciplinary Science Reviews* 48 (2):350–65.

Irwin, A. 2002. *Citizen Science: A Study of People, Expertise and Sustainable Development*. London: Routledge.

Irwin, A. 2006. “The Politics of Talk: Coming to Terms with the ‘New’ Scientific Governance.” *Social Studies of Science* 36 (2):299–320.

Isoaho, K., D. Burgas, N. Janasik, M. Mönkkönen, M. Peura, and J. Hukkinen. 2019. “Changing Forest Stakeholders’ Perception of Ecosystem Services with Linguistic Nudging.” *Ecosystem Services* 40:101028.

Jasanoff, S. 1997. *Science at the Bar: Law, Science, and Technology in America*. Cambridge, MA: Harvard University Press.

Jasanoff, S. 2003a. “Breaking the Waves in Science Studies: Comment on H. M. Collins and Robert Evans, ‘The Third Wave of Science Studies.’” *Social Studies of Science* 33 (3):389–400.

Jasanoff, S. 2003b. “(No?) Accounting for Expertise.” *Science and Public Policy* 30 (3):157–62.

Jasanoff, S. 2003c. “Technologies of Humility: Citizen Participation in Governing Science.” *Minerva* 41 (3):223–44.

John, S. 2018. “Epistemic Trust and the Ethics of Science Communication: Against Transparency, Openness, Sincerity and Honesty.” *Social Epistemology* 32 (2):75–87.

Johnson, B. B., and P. Slovic. 1995. “Presenting Uncertainty in Health Risk Assessment: Initial Studies of Its Effects on Risk Perception and Trust.” *Risk Analysis* 15 (4):485–94.

Jorm, A. 2025. *Expert Consensus in Science*. Singapore: Palgrave Macmillan.

Kitcher, P. 1990. “The Division of Cognitive Labor.” *Journal of Philosophy* 87:5–22.

Kitcher, P. 2011. *Science in a Democratic Society*. Amherst, NY: Prometheus Books.

Krimsky, S. 1984. “Epistemic Considerations on the Value of Folk-Wisdom in Science and Technology.” *Review of Policy Research* 3 (2):246–63.

Kuhn, T. S. 1996. *The Structure of Scientific Revolutions*. 3rd ed. Chicago: University of Chicago Press.

Lane, D. 2025. “Should You Defer to Individual Experts?” *Philosophy and Phenomenological Research* 111 (1):216–34.

Lavrakas, P. J. 2008. “Panel Fatigue.” P. 570 in Encyclopedia of Survey Research Methods. Edited by P. J. Lavrakas. Thousand Oaks, CA: Sage.

Lederman, N. G. 2013. “Nature of Science: Past, Present, and Future.” Pp. 831–79 in *Handbook of Research on Science Education*, edited by S. K. Abell and N. G. Lederman. New York: Routledge.

Levy, N. 2022. *Bad Beliefs*. Oxford: Oxford University Press.

Lewandowsky, S., G. E. Gignac, and S. Vaughan. 2013. “The Pivotal Role of Perceived Scientific Consensus in Acceptance of Science.” *Nature Climate Change* 3 (4):399–404.

Longino, H. E. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press.

Lynas, M., B. Z. Houlton, and S. Perry. 2021. “Greater than 99% Consensus on Human-Caused Climate Change in the

Peer-Reviewed Scientific Literature.” *Environmental Research Letters* 16 (11):114005.

Maertens, R., F. Anseel, and S. van der Linden. 2020. “Combatting Climate Change Misinformation: Evidence for Longevity of Inoculation and Consensus Messaging Effects.” *Journal of Environmental Psychology* 70:101455.

Mitchell Finnigan, S., J. Sheppard, and P. Vickers. 2023. *IASCPolls: The Institute for Ascertaining Scientific Consensus Polling Platform*. Zenodo. https://zenodo.org/records/8155054.

Morgan, M. G., and D. W. Keith. 1995. “Subjective Judgements by Climate Experts.” *Environmental Science & Technology* 29 (10):468A–476A.

Myers, K. F., P. T. Doran, J. Cook, J. E. Kotcher, and T. A. Myers. 2021. “Consensus Revisited: Quantifying Scientific Agreement on Climate Change and Climate Expertise among Earth Scientists 10 Years Later.” *Environmental Research Letters* 16 (10):104030.

Oreskes, N. 2004. “The Scientific Consensus on Climate Change.” *Science* 306 (5702):1686.

Oreskes, N. 2019. *Why Trust Science?* Princeton, NJ: Princeton University Press.

Oreskes, N., and E. M. Conway. 2010. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. New York: Bloomsbury Press.

Ottinger, G. 2013. *Refining Expertise: How Responsible Engineers Subvert Environmental Justice Challenges*. New York: New York University Press.

Pérez, E. O., and M. Tavits. 2017. “Language Shapes People’s Time Perspective and Support for Future-Oriented Policies.” *American Journal of Political Science* 61 (3):715–27.

Pidgeon, N., and B. Fischhoff. 2013. “The Role of Social and Decision Sciences in Communicating Uncertain Climate Risks.” Pp. 319–42 in Effective Risk Communication. Edited by J. Árvai and L. Rivers III. New York: Routledge.

Pielke, R. A., Jr. 2007. *The Honest Broker: Making Sense of Science in Policy and Politics*. Cambridge: Cambridge University Press.

Powell, J. 2019. “Scientists Reach 100% Consensus on Anthropogenic Global Warming.” *Bulletin of Science, Technology & Society* 37 (4):183–84.

Pöyhönen, S. 2017. “Value of Cognitive Diversity in Science.” *Synthese* 194 (11):4519–40.

Proctor, R. N., and L. Schiebinger. 2008. *Agnotology: The Making and Unmaking of Ignorance*. Stanford, CA: Stanford University Press.

Rabinovich, A., and T. A. Morton. 2012. “Unquestioned Answers or Unanswered Questions: Beliefs about Science Guide Responses to Uncertainty in Climate Change Risk Communication.” *Risk Analysis* 32 (6):992–1002.

Rawls, J. 1996. *Political Liberalism*. New York: Columbia University Press.

Rawls, J. 2001. *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press.

Rhodes, T., and K. Lancaster. 2022. “Uncomfortable Science: How Mathematical Models, and Consensus, Come to Be in Public Policy.” *Sociology of Health & Illness* 44 (9):1461–80.

Rip, A. 2003. “Constructing Expertise: In a Third Wave of Science Studies?” *Social Studies of Science* 33 (3):419–34.

Rode, J. B., M. Remshard, L. J. Groot, and S. van der Linden. 2025. “A Meta-Analytic Structural Equation Analysis of the

Gateway Belief Model: Highlighting Scientific Consensus Increases Support for Public Action on Climate Change." *Current Opinion in Behavioral Sciences* 63:101521.

Santana, C. 2021. "Let's Not Agree to Disagree: The Role of Strategic Disagreement in Science." *Synthese* 198 (25):6159–77.

Sharp, Z. 2009. *Annual Editions: Environment 09/10*. New York: McGraw-Hill.

Simpson, G. G. 1943. "Mammals and the Nature of Continents." *American Journal of Science* 241 (1):1–31.

Smolin, L. 2008. *The Trouble with Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next*. London: Penguin Books.

Solomon, M. 2006. "Groupthink versus the Wisdom of Crowds: The Social Epistemology of Deliberation and Dissent." *Southern Journal of Philosophy* 44 (S1):28–42.

Solomon, M. 2015. *Making Medical Knowledge*. Oxford: Oxford University Press.

Solomon, M. 2017. “Expert Consensus.” Pp. 353–60 in *The Routledge Companion to Philosophy of Medicine*. Edited by M. Solomon, J. R. Simon, and H. Kincaid. New York: Routledge.

Sturgis, P., and N. Allum. 2004. “Science in Society: Re-Evaluating the Deficit Model of Public Attitudes.” *Public Understanding of Science* 13 (1):55–74.

Suldovsky, B., and H. Akin. 2023. “The Role of Trust in Communicating Scientific Consensus and the Environmental Benefits of Genetically Engineered Crops: Experimental Evidence of a Backfire Effect.” *Environmental Communication* 17 (1):101–18.

Thoma, J. 2015. “The Epistemic Division of Labor Revisited.” *Philosophy of Science* 82 (3):454–72.

Thorp, H. H. 2025. “Convergence and Consensus.” *Science* 388 (6745):339.

Tosoni, S., and A. Ricotti. 2024. “Scientismo ingenuo e comunicazione della scienza: la lezione dell’infodemia.” Pp. 53–68 in Capire e comunicare la scienza: Conoscenze e scelte condivise in una società aperta. Edited by F. Giudice. Milan: Vita e Pensiero.

Towler, O. 2025. “Pedagogical Reforms in School Science Education for Mitigating Damaging Misinformation in a ‘Post-Truth’ World.” *Educational Theory* 75 (5):865–90.

van der Linden, S. L. 2021. “The Gateway Belief Model (GBM): A Review and Research Agenda for Communicating the Scientific Consensus on Climate Change.” *Current Opinion in Psychology* 42:7–12.

van der Linden, S., A. Leiserowitz, and E. Maibach. 2018. “Scientific Agreement Can Neutralize Politicization of Facts.” *Nature Human Behaviour* 2 (1):2–3.

Velez, Y. R., Á. Saavedra Cisneros, and J. Gomez. 2024. “Language Barriers: Causal Evidence of Linguistic Item Bias in Multilingual Surveys.” *Political Behavior* 46 (2):1213–32.

Vickers, P. 2023. *Identifying Future-Proof Science*. Oxford: Oxford University Press.

Vickers, P., L. Adamo, M. Alfano, C. Clark, E. Cresto, H. Cui, H. Dang, et al. 2024. “Development of a Novel Methodology for Ascertaining Scientific Opinion and Extent of Agreement.” *PLOS ONE* 19 (12):e0313541.

Vickers, P., C. Cowie, J. Hamilton, S. McMahon, and S. Mitchell Finnigan. 2026. “Comparing Astrobiologists’ Confidence in

Extraterrestrial Life Claims for K2-18 b and Cheyava Falls.” *Nature Astronomy* 10:774–76. https://doi.org/10.1038/s41550-026-02876-9.

Vickers, P., E. Gardiner, C. Gillen, B. Hyde, C. Jeancolas, S. Mitchell Finnigan, J. N. Nováková, et al. 2025. “Surveys of the Scientific Community on the Existence of Extraterrestrial Life.” *Nature Astronomy* 9 (1):16–18.

Weisberg, M., and R. Muldoon. 2009. “Epistemic Landscapes and the Division of Cognitive Labor.” *Philosophy of Science* 76 (2):225–52.

Winsberg, E., J. Brennan, and C. W. Surprenant. 2020. “How Government Leaders Violated Their Epistemic Duties during the SARS-CoV-2 Crisis.” *Kennedy Institute of Ethics Journal* 30 (3):215–42.

Wynne, B. 1982. *Rationality and Ritual: The Windscale Inquiry and Nuclear Decisions in Britain*. Chalfont St Giles, Bucks: British Society for the History of Science.

Wynne, B. 1992. “Misunderstood Misunderstanding: Social Identities and Public Uptake of Science.” *Public Understanding of Science* 1 (3):281–304.

Wynne, B. 2003. “Seasick on the Third Wave? Subverting the Hegemony of Propositionalism: Response to Collins & Evans (2002).” *Social Studies of Science* 33 (3):401–17.

Wynne, B. 2008. “Elephants in the Rooms Where Publics Encounter ‘Science’?: A Response to Darrin Durant, ‘Accounting for Expertise: Wynne and the Autonomy of the Lay Public.’” *Public Understanding of Science* 17 (1):21–33. https://doi.org/10.1177/0963662507085162.

Zehr, S. C. 2000. “Public Representations of Scientific Uncertainty about Global Climate Change.” *Public Understanding of Science* 9 (2):85–103.

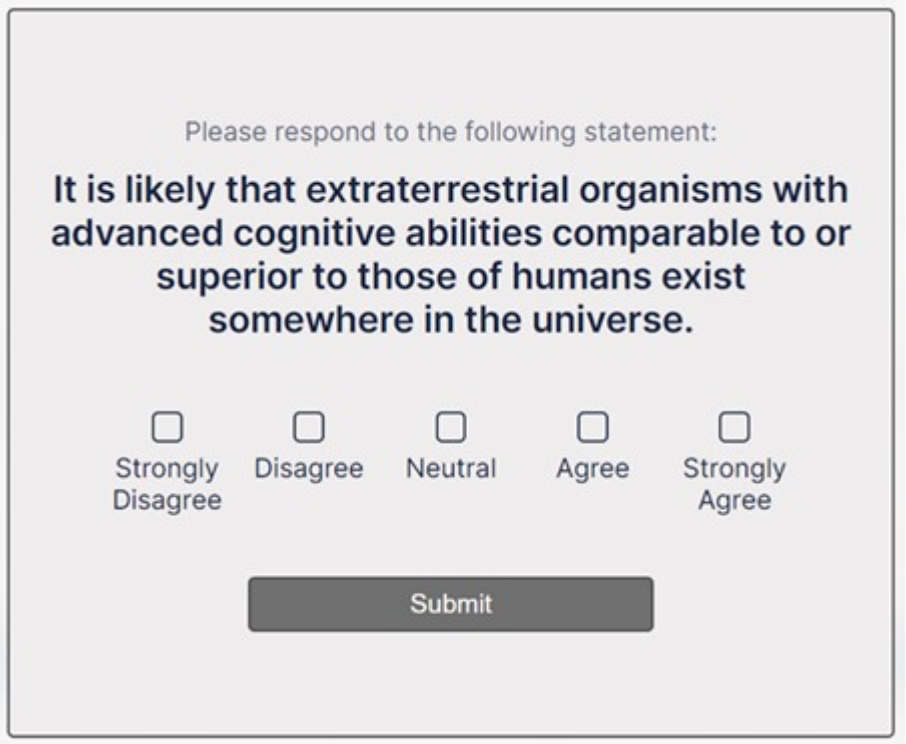


Figure 1. One of the surveys in the second pilot project.