Chapter 7. Self-Governance and Peer Review in Science-for-Policy: The Case of the IPCC Second Assessment Report[i]

In Clark Miller and Paul N. Edwards, eds., *Changing the Atmosphere: Expert Knowledge and Environmental Governance* (Cambridge, MA: MIT Press, 2001

by

Paul N. Edwards
Associate Professor
School of Information
University of Michigan

Stephen H. Schneider
Professor, Dept. of Biological Sciences
Senior Fellow, Institute for International Studies
Stanford University

In the spring of 1996, the Intergovernmental Panel on Climate Change (IPCC) released its long-awaited Second Assessment Report (SAR) on possible human impacts on the global climate system. The report's eighth chapter concluded that "the balance of evidence suggests that there is a discernible human influence on global climate" (Houghton et al. 1996, 5) — a phrase that has since become probably the single most-cited sentence in the IPCC's history. The Global Climate Coalition (an energy industry lobby group) and a number of "contrarian" scientists immediately launched a major, organized attack designed to discredit the report's conclusions, especially those relating to the crucial question of whether human activities are responsible for changes in the world's climate.

Led by the eminent physicist Frederick Seitz, these critics claimed that the IPCC had inappropriately altered a key chapter for political reasons. They alleged that the IPCC had "corrupted the peer review process" and violated its own procedural rules. These charges ignited a major debate, widely reported in the press, lasting several months.

The accusations of corruption reach a fundamental issue in the emerging global climate regime: namely, how the IPCC as a self-governing institution can maintain scientific integrity in the face of intense political pressures (both internal and external) and tightly constrained deadlines. In this chapter we consider these charges on three levels. First, we evaluate their accuracy as specific challenges to the IPCC peer review process, and note how the IPCC rules of procedure might be clarified to avoid them in the future. Second, we explore their meaning against the larger background of the IPCC's role in the politics of climate change. Finally, we use this episode to examine more fundamental questions about the role of formal review mechanisms in certifying scientific knowledge produced for policy contexts, and about the relative importance of those mechanisms in different national and cultural contexts.

The IPCC Second Assessment Report

The IPCC is an office of the United Nations Environment Programme and the World Meteorological Organization. Its purpose is to evaluate and synthesize the scientific understanding of global climate change for national governments and United Nations agencies, as expert advice for use in the ongoing negotiations under the Framework Convention on Climate Change (FCCC). The agency's nominal goal is to represent fairly the full range of credible scientific opinion. Where possible, it attempts to identify a consensus view on the most likely scenario(s). When consensus cannot be reached, the agency's charge is to summarize the major viewpoints and the reasons for disagreement. IPCC reports are intensively peer-reviewed. They are regarded by most scientists and political leaders as the single most authoritative source of information on climate change and its potential impacts on environment and society.

Like all IPCC assessments, the SAR contained three "Summaries for Policymakers" (SPMs), one for each of the IPCC's three Working Groups: climate science (Working Group I), impacts of climate change (Working Group II), and economic and social dimensions (Working Group III) (Bruce, Lee, and Haites 1996; Houghton et al. 1996; Watson et al. 1996). Since the full SAR stretches to well over 2,000 pages — most of it dense technical prose — few outside the scientific community are likely either to read it in its entirety or to understand most of its details. Therefore, these summaries tend to become the basis for press reports and public debate. For this reason, the Working Groups consider their exact wording with extreme care before they are published. At the end of the IPCC report process, they are approved word for word by national government representatives at a plenary meeting attended by only a fraction of the lead authors.

The SPM for Working Group I, which assesses the state of the art in the physical-science understanding of climate change, contained the following paragraph:

> Our ability to quantify the human influence on global climate is currently limited because the expected signal is still emerging from the noise of natural variability, and because there are uncertainties in key factors. These include the magnitude and patterns of long–term natural variability and the time–evolving pattern of forcing by, and response to, changes in concentrations of greenhouse gases and aerosols, and land surface changes. Nevertheless, the balance of evidence suggests that there is a discernible human influence on global climate (Houghton et al. 1996, 5, emphasis added).

Three-quarters of this paragraph consists of caveats about uncertainties and limitations of current understanding. Nonetheless, its now-famous closing sentence marked the first time the IPCC had reached a consensus on two key points: first, that global warming is probably occurring ("detection"), and second, that human activity is more likely than not a significant cause ("attribution"). Like this summary paragraph, the body of the report discussed — frequently and at length — the large scientific uncertainties about attribution. The Working Group carefully crafted the SPM's "balance of evidence" sentence to communicate the strong majority opinion that despite these uncertainties, studies were beginning to converge on a definitive answer to the attribution question.

The SAR was fraught with political significance. Official publication of the full report occurred in early June, 1996. At that point the Second Conference of Parties to the FCCC (COP-2) was about to meet in Geneva; the session would determine some of the starting points for the Kyoto meeting in 1997, where binding greenhouse-gas emissions targets and timetables were to be negotiated. A sea change in American climate policy was widely rumored. Since the Reagan administration, official US policy had sanctioned only voluntary, non-binding emissions targets and further scientific research. If the United States were to abandon its resistance to binding emissions targets and timetables, a strong international greenhouse policy would become much more likely. Since the more-research, no-binding-targets position was officially based on assertions that scientific uncertainty remained too high to justify regulatory action, the SAR's expressions of increased scientific confidence were viewed as critical.

The rumors proved correct. On July 17, 1996, then US Under-Secretary of State for Global Affairs Tim Wirth formally announced to COP-2 that the United States would now support "the adoption of a realistic but binding target" for emissions. The exact degree to which the IPCC SAR influenced this policy change cannot be known. But Wirth certainly gave the impression that the report was its proximate cause. He noted in his address that "the United States takes very seriously the IPCC's recently issued Second Assessment Report." He then proceeded to quote the SAR at length, proclaiming that "the science is convincing; concern about global warming is real" (Wirth 1996, emphasis added).

"A Major Deception on Global Warming"

On June 12, 1996, just days after formal release of the IPCC SAR and scant weeks before the COP-2 meeting in Geneva, the Wall Street Journal (WSJ) published an op-ed piece entitled "A Major Deception on Global Warming." The article was written by Frederick Seitz, President Emeritus of Rockefeller University. Seitz is not a climate scientist but a physicist. Nevertheless, his scientific credentials are formidable. He is a recipient of the National Medal of Science and a past President of both the National Academy of Sciences and the American Physical Society.

In his article, Seitz accused some IPCC scientists of the most "disturbing corruption of the peer-review process" he had ever witnessed (Seitz 1996).

Seitz's Accusations

Seitz's distress stemmed from the fact that the lead authors of the SAR's Chapter 8 — on detection and attribution — had altered some of its text after the November, 1995 plenary meeting of Working Group I (WGI), in Madrid, at which time the chapter was formally "accepted" by the Working Group. According to Seitz, since the scientists and national governments who accepted Chapter 8 were never given the chance to review the truly final version, these changes amounted to deliberate fraud and "corruption of the peer-review process." Not only did this violate normal peer review procedure, Seitz charged; it also violated the IPCC's own procedural rules.

Quoting several sentences deleted from the final version of the chapter, Seitz argued that the changes and deletions "remove[d] hints of the skepticism with which many

scientists regard claims that human activities are having a major impact on climate in general and on global warming in particular." Without directly attributing motives, Seitz implied that the changes had been made in the interests of promoting a particular political agenda. Seitz said that Benjamin D. Santer, lead author of Chapter 8, would have to shoulder the responsibility for the "unauthorized" changes. Seitz was not present at the IPCC meetings. He did not contact Santer or anyone else at the IPCC to verify that the changes were indeed "unauthorized" before publishing his op-ed piece.

Responses from Santer and the IPCC

Santer responded immediately, in a letter co-signed by some 40 other IPCC officials and scientists (myself among them — SHS). They said that Seitz had misinterpreted the IPCC rules of procedure. Rather than being "unauthorized," they wrote, the post-Madrid changes were in fact required by IPCC rules, under which authors must respond to comments submitted during peer review or arising from discussions at the meetings (Santer et al. 1996a).

Commentators at the Madrid meeting had advised making changes to Chapter 8 for two reasons. First, they urged clarification of the meaning and scientific content of some passages in accordance with the recommendations of reviewers (including some criticisms introduced at the Madrid meeting itself). Second, they thought the structure of the chapter should be brought into conformity with that of other SAR chapters. In particular, a "Concluding Summary" was removed from the final version, since no other chapter contained a similar section. (Chapter 8, like all the rest, already had an "Executive Summary.") Sir John Houghton, in his capacity as co-chairman of WGI, specifically authorized that these changes be made, though he did not review their wording.

Santer, in consultation with other Chapter 8 authors, made the suggested changes in early December. The entire SAR, including the newly revised Chapter 8, was "accepted" by the full IPCC Plenary at Rome later that month.

Santer made the changes himself, and the final version of the chapter was not reviewed again by others. However, as he and his colleagues continually stressed, this procedure was the normal and agreed IPCC process. Santer et al. pointed out that no one within the IPCC objected (or had ever objected) to this way of handling things. Replying separately in support of Santer and his colleagues, IPCC Chairman Bert Bolin and WGI Co-Chairmen John Houghton and L. Gylvan Meira Filho quoted the official US government review of Chapter 8, which stated explicitly that "it is essential that... the chapter authors be prevailed upon to modify their text in an appropriate manner following discussion in Madrid" (Bolin, Houghton, and Meira-Filho 1996)

Further Exchanges

The Wall Street Journal op-ed was not the first time charges of suppression of scientific uncertainty in Chapter 8 had been aired. On May 22, a few days before the Seitz op-ed appeared, the small journal Energy Daily reported the same allegations in considerably greater detail (Wamsted 1996). The Energy Daily article also reported their source: a widely circulated press release of the Global Climate Coalition (GCC, an energy industry lobby group).

In its June 13 issue, the prestigious scientific journal Nature also reported on the GCC allegations (Masood 1996). The Nature report, unlike the Seitz and Energy Daily articles, included explanations of the revision and review process from Santer and the IPCC. Under the hot-button headline "Climate report 'subject to scientific cleansing,'" an accompanying editorial argued that the GCC analysis was politically motivated and generally false. But the editorial also noted that the Chapter 8 changes may have resulted "in a subtle shift... that... tended to favour arguments that aligned with [the SAR's] broad conclusions" (Nature editors 1996).

The Wall Street Journal op-ed set off a lengthy chain of exchanges lasting several months. The main participants in the public controversy were Seitz, Santer, other Chapter 8 authors, the Chairmen of the IPCC (Sir John Houghton and Bert Bolin), and climate-change skeptics S. Fred Singer and Hugh Ellsaesser. Singer, in particular, made the charges of political motivation explicit. In a letter to the Wall Street Journal, he wrote that Chapter 8 had been "tampered with for political purposes." The IPCC, he claimed, was engaged in a "crusade to provide a scientific cover for political action" (Singer 1996).

Semi-privately, in electronic mail exchanges involving many additional participants (and widely copied to others), the debate became intense and sometimes quite bitter. Santer, who felt forced to defend himself, spent the majority of his summer time responding to the charges. Previously a quiet, private man known to scientists primarily as a proponent of the rigorous use of statistical methods, Santer rapidly became a public figure, submitting to dozens of interviews. The drain on his time and energy during this period kept him from his scientific work, he said (personal communication, 1996).

Both the public and the private exchanges themselves became objects of further press reports, widely disseminated by the news wire services. As they went on, the debate spread from the initial issues about peer review and IPCC procedure to include questions about the validity of Chapter 8's scientific conclusions. Even before the report was formally published, climate-change skeptics had claimed that Chapter 8 dismissed or ignored important scientific results that disconfirmed the global warming hypothesis. They argued that the allegedly illegitimate changes to Chapter 8 made this problem even more acute (Brown 1996).

The Chapter 8 Revisions and IPCC Self-Governance

As a hybrid science-policy body, the IPCC must maintain credibility and trust vis-à-vis two rather different communities: the scientists who make up its primary membership, and the global climate policy community to which it provides input. Independent self-governance is one of the primary mechanisms by which it achieves this goal. The IPCC's rules of procedure spell out a variety of methods designed to ensure that its reports include the best available scientific knowledge and that they represent this knowledge fairly and accurately. Chief among these is the principle of peer review, traditionally one of the most important safeguards against bias and error in science.

Seitz, the GCC, and others accused the authors of Chapter 8 of fraud on two counts. First, they alleged that the changes made to Chapter 8 after the final IPCC plenary violated the IPCC's own rules of procedure. Second, and more seriously, they charged

them with violating the fundamental standards of scientific peer review. In this section, we argue that IPCC rules were not violated in the case of Chapter 8. In addition, we argue that in practice the process correctly reflects the essential tenets of peer review. However, we also show that the IPCC rules do not specify adequate closure mechanisms for the report drafting process. We demonstrate that the two-level certification process ("acceptance" and "approval" of IPCC documents) is poorly specified as well, and can even invite misinterpretation by determined critics.

In their responses to the Seitz/GCC charges, the Chapter 8 authors claimed that IPCC governance rules required them to make the changes advised immediately before and during the Madrid WGI Plenary. Analysis of the IPCC rules suggests that the real situation is more ambiguous. Yet they had three very good reasons for believing this to be the case.

First, the rules require authors to respond to commentary, to the best of their ability and as fully as possible (Intergovernmental Panel on Climate Change 1993). Working Group co-chairs have broad discretion to define this process and set time limits for it. Nowhere do IPCC rules explicitly address the question of when a report chapter becomes final (i.e., when all changes must cease). Therefore, Santer et al. correctly understood that the Working Group Chairs and the Plenary meeting itself would define the endpoint of the revision process.

Second, report chapters are "accepted" rather than "approved." Acceptance constitutes IPCC certification that the drafting and review process has been successfully completed. It is an expression of trust in the authors and the process, and is explicitly distinguished from "approval," or detailed review on a line-by-line basis. Operating under these definitions, the IPCC Plenary "approved" the WGI Summary for Policymakers (SPM), but "accepted" Chapter 8. In other words, Plenary acceptance did not imply word-for-word review of the chapter. Instead, it indicated trust that the authors had responded appropriately and sufficiently to the review process. Therefore, the Chapter 8 authors believed that the rules permitted them to make changes when explicitly requested to do so by the IPCC Plenary, or in response to peer comments received at or immediately prior to the Plenary.

Third, no IPCC member nation ever seconded the Seitz/GCC objections (Bolin 1996). (Ninety-six countries were represented at the Madrid plenary.) From this, above all, we can safely infer that Santer et al. proceeded exactly as expected. They believed that they were following IPCC rules, and this made perfect sense within the established informal culture of the IPCC.

However, a careful reading of the IPCC's formal rules reveals that in fact the rules neither allow nor prohibit changes to a report after its formal acceptance. The legalistic Seitz/GCC reading of the rules is not, therefore, completely implausible — even if it was, as we believe, primarily a smokescreen to divert attention from the clear consensus that attribution could no longer be considered unlikely.

Our analysis suggests a significant flaw in the rules as currently written. While "approved" documents (the SPMs) clearly must not be altered once approved, there is no precisely defined closure mechanism for "accepted" documents (full-length Working Group reports and their constituent chapters) (Intergovernmental Panel on Climate Change 1993). The Seitz/GCC attack has effectively demonstrated that a hybrid

science/policy organization like the IPCC needs better, more explicit rules of procedure. This minor virtue aside, however, the Seitz/GCC reading violates the spirit and intent of the IPCC process.

The IPCC is run by scientists. Its participants think of it primarily as a scientific body. By the standards of many political organizations, its formal rules of governance are not very extensive. They are also not very specific. The rules purposely leave undefined the meaning of key terms such as “expert” and important processes such as “taking into account” comments. Under the rules, Lead Authors carry full responsibility for report chapters, and the IPCC leadership retains very broad discretion, subject to Plenary “acceptance” and “approval” by national governments.

There are good reasons for this arrangement. Formal governance is relatively unimportant in scientific culture. This is true because scientists generally belong to small social groups endowed with strong and deeply entrenched (informal) norms. In addition, since scientific methods and results are constantly changing, too much focus on formal rules would inhibit progress. Likewise, formal rules are not very important in the day-to-day functioning of the IPCC. Instead, informal rules based on the everyday practices of scientific communities guide the bulk of the work (Collins and Pinch 1993; Gilbert and Mulkay 1984; Latour and Woolgar 1979; Merton 1973).

Maintaining this informality is quite important for effective scientific work. Yet it is not without dangers, especially in a situation where almost any scientific finding can have political implications (Jasanoff 1990; Jasanoff and Wynne 1998). Just as in any other politicized realm, without clear procedures to ensure openness and full rights of participation, dissenters may find — or believe they have found — their voices ignored. One of the IPCC’s most important features is its openness and inclusivity; balancing this against scientific informality will require constant vigilance, and perhaps a reconsideration of the formal review process.

From the point of view of political legitimacy, then, acceptance of reports before final revision is clearly a risky proposition (Jasanoff 1991). But from the viewpoint of scientific legitimacy, ongoing revision is a normal feature of the research cycle. Even after a multi-stage review process, minor flaws can be found and improvements added. This is not unlike the common situation in which an author makes minor changes to the galley proofs of a manuscript — changes not subject to peer review. Thus, in the case of the IPCC, adding a final approval stage to the already long and cumbersome review process would be unlikely to add significantly to the scientific credibility of the final result. While it needs to revise its rules to better protect itself from accusations of political capture, the IPCC must also, at all costs, avoid becoming a science-stifling, inflexible bureaucracy.

In fact, in late 1999 the IPCC finalized a major revision to its rules of procedure, in response to considerations that included the Chapter 8 controversy. According to David Griggs of the IPCC Working Group I Technical Support Unit, one of the major changes is the introduction of “Review Editors.” These editors

> will assist the Working Group Bureaux in identifying reviewers for the expert review process, ensure that all substantive expert and government review comments are afforded appropriate consideration, advise Lead Authors on how to handle contentious/controversial issues and ensure

genuine controversies are reflected adequately in the text (Griggs, personal communication, July 1999).

The new rules should make disputes such as the Chapter 8 controversy less frequent and, perhaps, provide new mechanisms for resolving them without resort to salvos in the popular press — although the already-adequate existing mechanisms did not prevent Seitz and his colleagues from sidestepping them in order to attack the IPCC.[ii]

The Chapter 8 Revisions and the Peer Review Process

As we noted above, one of the most important standards of scientific accountability holds that publications must be reviewed by expert peers before results are released. Seitz and the GCC accused the IPCC of violating this standard, too. Were they right?

The Peer Review Process

Peer review is among the oldest certification practices in science, established with the Philosophical Transactions of the Royal Society in 1655 (Chubin and Hackett 1990, 19).

In a typical peer review procedure, scientists write articles and submit them to a journal. The journal editor sends the article to several referees, all of them experts in the authors' field ("peers"). Peer review at journals is usually "blind": authors are not informed of the referees' identity, though the author's name may be known to the referees. Blind review operates on the principle that free expression of criticism is more likely when referees, who often know authors personally and want to maintain good relations with them, can say what they think without having to consider authors' reactions, especially to negative evaluations. Many journals use a more stringent "double-blind" procedure, in which neither referees nor author(s) are informed of each other's identity. Double-blind review is based on the principle that criticism is more impartial when authors' identities are unknown to referees, who might be swayed in either positive or negative directions by authors' reputations, personality traits, etc. A similar process is normally applied to grant proposals (Kassirer and Campion 1994). Standard peer review procedure varies by field and by journal or grant agency. Few journals in the atmospheric sciences, for example, use double-blind review.

However, the fundamental purpose of peer review is to strengthen the quality of work by subjecting it to criticism and evaluation by those best qualified to judge it. Many paths can lead to this goal; it does not depend on blind procedures, and not all journals employ them. Some journal editors (like myself, as editor of Climatic Change — SHS) go so far as to encourage referees to reveal themselves. In any case, many scientific communities are small enough that even double-blind referees and authors can often guess each other's identity.

Referees can typically choose one of three recommendations: acceptance, rejection, or acceptance after certain specified changes are made ("revise and resubmit"). The last of these responses is by far the most common. The authors then rewrite their article in response to the reviewers, and the editor serves as referee on the issue of whether the revisions have satisfactorily answered reviewers' criticisms. The process usually goes back and forth several times, with several rounds of revision, until a suitable compromise

is achieved among reviewers, authors, and the editor. (At grant agencies, grants officers fill a role analogous to the journal editor. The process usually requires resubmission of the grant application in subsequent rounds of the funding cycle.)

Does Peer Review Work — and For What?

Certainly peer review is imperfect. Not all referees do their job well, and personal, political, and social factors can all enter into the process in unseemly ways. Empirical evaluation of peer review's effectiveness is decidedly mixed.

But as we will show, the question of whether peer review "works" depends largely on what one thinks peer review is for. A science-studies approach to peer review gives a new perspective on its purpose, one more consonant with the high esteem in which scientists hold the process and with its role in IPCC assessments. First, however, let us briefly review the major criticisms of the process.

Numerous studies of peer review have judged that it fails as a dependable indicator of research quality. Several experiments have shown that agreement between referees on the same article is generally only slightly better than chance (for summaries of these see Cichetti 1991; Marsh and Ball 1989). A major, long-term study of grant proposal review at the National Science Foundation concluded that "funding of a specific proposal… is to a significant extent dependent on the applicant's 'luck' in the program director's choice of reviewers" (Cole 1992, 99). Other studies indicate that peer review suffers from systematic "confirmatory bias," i.e. the tendency to rate more highly studies which confirm existing beliefs, regardless of their quality (Bornstein 1991; Cole 1992; Mahoney 1977; Ross 1980). A related critique views peer review as a form of censorship which effectively blocks expression of innovative ideas that challenge dominant scientific paradigms (Moran 1998). Finally, peer review cannot reliably detect fraudulent science (Chubin and Hackett 1990, Chapter 5).

Several scholars have claimed that far from assuring impartiality, "blind" peer review in fact encourages two kinds of counterproductive, unethical referee behavior. First, anonymous referees — whose rewards for their efforts are minor to non-existent — may tend to minimize the time they devote to review, even to the point of approving work they have not actually read. Second, anonymity allows referees to engage more easily in personally or politically motivated attacks on others' work; as we pointed out above, even double-blind review may not prevent authors and referees from guessing each other's identities in a small field (and wrong guesses can turn out to be even more harmful than right ones). One major study of peer review recommended eliminating the "blind" system by requiring referees to sign their reviews:

> This would hold reviewers publicly accountable for their decisions and would take a step toward acknowledging the value of reviewers' work. No longer would it be convenient for a reviewer to trash another's work. Nor would it be advisable to endorse unexamined work (Chubin and Hackett 1990).

Some commentators have recommended abolishing the system altogether (Roy 1985). Others see it as withering away of its own accord under the influence of reviewer fatigue, ever-expanding numbers of publications, and new electronic media that can circumvent the process, either deliberately or not (Judson 1994).

A Revised Conception of Peer Review

Most of these criticisms of peer review depend on a particular (and often tacit) view of its purpose, namely that peer review acts as a kind of "truth machine," automatically separating "good" science from "bad." This view — which may be more common among those studying peer review than among participants in the process — implicitly assumes that scientists (peers) agree very closely about most things, so that the opinion of one scientist about an article ought to be similar to that of most others. That this should turn out, empirically, not to be the case is surprising only if one subscribes to what Stephen Cole, author of the NSF peer review study mentioned above, calls "the mythology that scientists do not or should not disagree." He cites the University of Chicago statistician William Kruskal, Jr.: "…careful objective studies of expert judgment typically find them disagreeing more than had been expected and more than the experts themselves find comfortable. …Variability is usually considerable, the more so in close cases (Kruskal, cited in Cole 1992, 100)."

Yet disagreement among peers is undesirable only if it is interpreted (wrongly, we would argue) as arbitrary. Taking a science-studies view of the process, Cole goes on to observe that disagreement among experts is basic to scientific practice.

> The great majority of reviewer disagreement observed [in our empirical studies] is probably a result of real and legitimate differences of opinion among experts about what good science is or should be. …Contrary to a widely held belief that science is characterized by agreement about what is good work, who is doing good work, and what are promising lines of inquiry, this research indicates that concerning work at the research frontier there is substantial disagreement in all scientific fields (Cole 1992, 100).

In fact, disagreement is vital to science, since it drives further investigation. But if expert judgment varies too widely to provide a quasi-mechanical means of winnowing out bad science from good, why is peer review important? The answer depends on one's conception of its role and purpose.

We maintain that peer review ought to be regarded as a human process whose primary functions are to improve the quality of scientific work, to maintain accountability both inside and outside the scientific community, and to build a scientific community that shares core principles and beliefs even when it does not agree in detail (Haas 1990a, 1990b). For example, peer review helps to minimize errors; reviewers frequently catch mathematical and methodological mistakes. Reviewers frequently also suggest better methods, recalculate numbers, and offer solutions to unresolved problems. Peer review also helps to distribute new research results, and helps assure fair distribution of credit for work done. It acts as a certification mechanism, a barrier to entry, and a disciplinary device (in many senses) (Foucault 1977; Kuhn 1962; Merton 1973). Despite the severe sound of these latter functions, they are in fact vital to building any coherent knowledge community.

Peer review can be also described as an institutionalized form of the "virtual witnessing" process by which science establishes factual knowledge (Shapin and Shaffer 1985). It ensures that at least a few relatively disinterested parties have carefully scrutinized the experimental procedure and the reasoning and agreed with the conclusions drawn by

the author(s).[iii] It is a form of accountability, a way for the community to rehearse (and enforce) its fundamental norms and practices. For some or all of these reasons, nearly every scientist regards peer review as an extremely important mechanism, even though most are aware of its problems.

Several empirical studies have reached positive verdicts on the ability of the process to improve the quality of publications despite its acknowledged failings (Abelson 1980; Daniel 1993). Under this conception, as long as fundamental standards of scientific practice are met, the purpose of peer review is to minimize disagreements, but not necessarily to resolve them — since disagreement is viewed as a natural and unavoidable element of science as a human practice. Such a concept is, in fact, explicitly recognized in the IPCC rules of procedure, which specify that

> …lead authors should clearly identify disparities of view for which there is significant scientific or technical support, together with the relevant arguments… It is important that reports describe different (possibly controversial) scientific or technical views on a subject, particularly if they are relevant to the political debate (Intergovernmental Panel on Climate Change 1993).

Finally, and most significantly for our purposes here, peer review plays a major role in establishing the credibility of expert knowledge for policy purposes (see Chapter 1, this volume). The power of "virtual witnessing" stems from its basic (and basically democratic) tenet that any suitably qualified person could (at least in principle) play the role of witness. Symbolically, if not literally, it establishes the openness of science to the whole human community. Echoing this point, Chubin and Hackett call peer review the "flywheel of science, if for no other reason than that it symbolizes the professional autonomy and the accountability of science to the society that sustains it. Peer review communicates and enforces the terms of a social contract (Chubin and Hackett 1990, 216)." As a "flywheel" of accountability, peer review dampens the influences of personal, social, and political interests that might otherwise affect science-for-policy. It also renders publication of both "junk science" and true paradigm-challenging innovation considerably more difficult.

Thus we believe that while its problems should not be ignored, its virtues must be recognized. Peer review must stand as a basic norm of scientific practice. But its purpose should be clarified. It is not a truth machine, but a human technique for quality improvement, accountability, and community building. With this conception of peer review in mind, we can now return to the Chapter 8 controversy, where both sides accepted peer review as a fundamental standard. The only question was whether the IPCC authors had attempted illegitimately to circumvent it.

Did the Chapter 8 Revisions "Corrupt" the Peer Review Process?

The first thing to note is that IPCC reports are not primary science, but assessments of the state of the field. In other words, they do not constitute new research, but analysis and evaluation of existing, previously peer-reviewed research. IPCC authors sometimes incorporate research that has not yet been reviewed or published. In such cases, manuscripts must be made available to both IPCC authors and reviewers. This often means that peer review is already in process, but not yet completed. It also allows the IPCC to consider non-peer reviewed, but potentially valuable sources such as "articles

published in industry or trade journals; proceedings of workshops; reports and working papers of research institutions, private firms, government agencies, and non-government organizations; contractor reports prepared for government agencies, firms, industry groups, and other non-governmental organizations; and books that have not been peer reviewed (Leary 1999)." Such sources must be specifically flagged as "not peer reviewed," and IPCC authors are routinely directed to

> critically assess [their] quality and validity… Don't just cite results from non-peer reviewed sources without assessing their quality and validity. (Actually, you should be doing this for peer reviewed sources as well — our job is to assess the state of knowledge, not just report what's in the literature). Basically, your expertise substitutes for the peer review process for material that has not been peer reviewed (Leary 1999).

Indeed, IPCC rules specify that assessment report authors must rely upon "the peer-reviewed and internationally-available literature, including scientific and technical publications prepared by national governments and scientific bodies [and] the latest reports from researchers that can be made available in preprint form for IPCC review," in addition to IPCC-prepared supporting materials (Intergovernmental Panel on Climate Change 1993).

Nevertheless, as syntheses and evaluations IPCC reports must be (and are) subjected to their own peer review process, partially described above. The IPCC peer review procedure is far more extensive and inclusive than most. Most IPCC members (including non-specialists, such as governments and lobby groups) receive draft IPCC documents and may submit "peer" comments. IPCC rules specify that draft chapters be circulated to

- specialists who have significant publications in particular areas;
- lead authors, contributors, and reviewers on the IPCC lists maintained by the Working Group and Subgroup co-chairs;
- IPCC participating countries and organizations;
- specialist reviewers nominated by appropriate international scientific and technical organizations (e.g., WMO, UNEP, ICSU, Third World Academy of Sciences, FAO, IOC, World Bank, Regional Development Banks, OECD) (Intergovernmental Panel on Climate Change 1993).

Chapter authors are required to "take into account" all comments, although the meaning of this phrase is deliberately left vague. Given the volume of commentary and the many duplicate and irrelevant comments received, responses may be no more than a couple of words. Yet in aggregate, this extremely extensive peer review process typically leads to hundreds or even thousands of changes, as each document typically goes through several drafts.

Some of the most outspoken global-warming skeptics in fact participated in the formal peer review of the SAR, including Chapter 8. Since Seitz is not a climate scientist, and Singer is no longer active in research, they did not qualify as formal reviewers. However, Singer regularly attends IPCC meetings. In 1995, as the IPCC prepared the SAR, Singer was present at both the Madrid meeting and the IPCC plenary at Rome. Representatives from a number of NGOs which typically take a skeptical stance, including the Global Climate Coalition and several energy and automotive industry lobbies, also participated in the SAR peer review. Other skeptic referees included Patrick Michaels, Hugh

Ellsaesser, and MIT meteorologist Richard Lindzen, an outspoken critic of some aspects of climate modeling. Lindzen was recently appointed a Lead Author for the next IPCC Assessment Report. Thus the skeptical views of the Chapter 8 critics were already very well represented in the SAR peer review process, all the way through Chapter 8's formal acceptance at Rome.

Did the Chapter 8 authors "corrupt" the IPCC peer review process? Let's look at how it worked in IPCC Working Group I (WGI). In July of 1995, the third installment of the WGI drafting and review process for the SAR took place in Asheville, North Carolina. This meeting, like all other IPCC processes, was characterized by exceptional openness to critique, review, and revision. About six dozen climate scientists from dozens of countries took part. The meeting was designed to make explicit the points of agreement and difference among the scientists over exceedingly controversial and difficult issues, including Chapter 8 — the most controversial.

New lines of evidence had been brought to bear by three climate modeling groups around the world, each suggesting a much stronger possibility that a climate change signal has been observed and that its pattern (or fingerprint) is matched to anthropogenic changes. Ben Santer, as a Convening Lead Author of Chapter 8, had assembled the results of a number of modeling groups. He presented the results of his group's effort not just to Chapter 8's Lead Authors and contributors, as is typical in IPCC meetings, but to the entire scientific group assembled at Asheville.

In this setting, Santer had to explain this work not only to his most knowledgeable peers, but also to scores of others from diverse scientific communities. These included stratospheric ozone experts such as Susan Solomon and Dan Albritton, satellite meteorologists such as John Christy, and biosphere dynamics experts such as Jerry Melillo. Climatologists such as Tom Karl and I (SHS) were also present, along with the heads of national weather services and other officials from several countries who served on the IPCC's assessment team.

Not everybody present was equally knowledgeable on the technical details of the debate, of course. Perhaps only 25 percent of those assembled had truly in-depth knowledge of the full range of details being discussed. However, all understood the basic scientific issues and most knew how to recognize slipshod work -- to say nothing of a fraud or a "scientific cleansing" -- when they saw it. Even the less familiar participants thus served an essential role: they acted as technically-skilled witnesses to the process of open debate.

This remarkable session lasted for hours. (In fact, it was continued less formally after dinner by roughly a dozen scientists, who spent nearly three hours discussing the final paragraph of the "Detection Section" of the Summary for Policymakers. Ben Santer, personal communication.) Though occasionally intense, it was always cordial, never polemical. As a result, the wording of Chapter 8 was changed. Ideas and concepts were somewhat altered, but basic conclusions by and large remained unchanged — because the vast bulk of those assembled were convinced that the carefully hedged statements the lead authors proposed were, in fact, an accurate reflection of the state of the science based upon all available knowledge, including the new results.

This was peer review at ten times the normal level of scrutiny! It would be almost inconceivable for the editor of a peer-reviewed journal to duplicate this process. A few

referees and an editor can only hope to execute the reviewing role a fraction as well as the remarkable, open process at Asheville. Moreover, after the Asheville meeting, two more IPCC drafts were written and reviewed by hundreds of additional scientists from all over the globe.

It is true that the Asheville meeting was not a "blind" review, since everyone was in the same room. Under these circumstances, reputations, personalities, institutional politics, and the simple fatigue induced by long meetings probably played some role, one that might have been reduced through a more formal procedure where authors and respondents were more distanced from each other. Yet the Asheville meeting was only one part of a much more extensive process that did include the formal review described above. Furthermore, as we pointed out above, "blind" procedures are only one way to achieve the fundamental goals of peer review — and not necessarily the most effective one. Our claim here is that the quality improvements generated by the extensive and inclusive IPCC peer review process far outweigh the disadvantages of the open-meeting format for the final stages of peer review. If the real purposes of peer review are, as we have argued, quality improvement, accountability, and community-building, then the IPCC process is as near to an ideal example as it may be possible to find.[iv]

Furthermore, science-for-policy (such as the IPCC assessment reports) operates under severe time constraints not present in pure research. Reviewing and assessing the entirety of the rapidly growing climate-related scientific literature in two to three years is a vast project. Even full-time professional assessors would find this challenging, let alone the volunteer members of the IPCC. The IPCC's attempts to include a very wide range of interested parties in the review process greatly increases the work involved in responding to peer commentary. If the IPCC assessment reports are to serve their key function as input to FCCC climate negotiations, at some point the review-and-improvement cycle must stop.

An Open Process of Scientific Debate: Witnessing in Action

At Madrid, Santer presented Chapter 8's conclusions to the national delegates of 96 IPCC member nations. The conclusions were not presented alone, but followed a presentation to the plenary session of the scientific evidence contained in Chapter 8. Nevertheless, several countries objected to the Chapter 8 conclusions. Most of the objections came from OPEC or a few less-developed nations. One delegate, from Kenya, moved to have the chapter entirely dropped from the final report.

In response, the meeting's chair — following procedures often used at IPCC Plenary meetings to resolve disputes — called for a drafting group to revise the detection and attribution section of the Summary for Policymakers and to inform the Chapter 8 lead authors of various delegates' concerns. Nations complaining about the Chapter 8 draft were invited, indeed expected, to meet with Lead Authors, first to hear the scientists' point of view and then to fashion new, mutually acceptable language.

This breakout group worked for the better part of a day. Delegates from over half a dozen countries — including the Kenyan who had publicly advocated dropping the chapter — met with about half a dozen Chapter 8 authors, including Santer, co-Lead Author Tom Wigley, and scientists Kevin Trenberth, Michael MacCracken, John Mitchell, and me (SHS). The Kenyan sat next to me. Initially, he was confused by the discussion and somewhat hostile. We had many side conversations about what was being

discussed: models, data, statistical tests and various climate forcing scenarios. Although he was not a front-rank climate researcher, this delegate was a trained scientist. He began to grasp the nature of the Lead Authors' arguments, listening carefully to about half of the breakout meeting.

Ironically, the Saudi Arabian delegation sent no representative to this most controversial drafting group, even though Saudi Arabia had led the opposition in the plenary meeting. During the Chapter 8 debate, Saudi delegates often issued objections soon after receiving notes from the Global Climate Coalition representative. (Non-governmental organizations were also represented at Madrid. For example, Singer — President of the Science & Environmental Policy Project and a self-proclaimed skeptic — raised a number of issues from the floor.)

Later in the plenary meeting, when Santer presented the drafting group's revised text, the Saudi delegates once again objected. Santer forcefully challenged them. Why, he asked, had no Saudi attended the breakout group — if their objections had some basis in science? The head Saudi delegate haughtily announced that he didn't have to account to a lead author for his decisions about which drafting group to attend. Besides, he said, his was "only a small delegation" of a few people.

At this point the Kenyan delegate rose to speak. "I'm a member of a small delegation too," he said. (He was the only Kenyan representative.) "But somehow I managed to attend this most important drafting session. As a result, I am convinced that Chapter 8 is now well written and I have no objections to its inclusion in the report." (A paraphrase of his words from memory, by SHS.) The impact of his intervention was stunning, stopping with a few words what appeared to be a mounting movement of OPEC and LDC opposition to Chapter 8 before it could garner any further support.

Later on I (SHS) privately congratulated the Kenyan for having the courage to object publicly, observe privately, and then re-evaluate his position before the entire plenary. He said he wasn't sure his country would approve of his stance, but having witnessed the debate process for several hours, he had become convinced it was honest and open. That was all he needed to change his opinion from preconceived skepticism to support of the Lead Authors' conclusions.

What this courageous delegate did was the essence of good science. He allowed his initial hypothesis to be subjected to new evidence, tested it, and found it wanting. He then listened to arguments for a different point of view, subjected them to the tests of evidence and debate, and reached a new conclusion.

Contrast this open IPCC process with that of the critics led by Seitz and the Global Climate Coalition. The latter first presented their technical counter-arguments in the editorial pages of the Wall Street Journal. Some alleged — falsely, and without evidence — that Chapter 8's conclusions were based upon non-peer-reviewed articles (Santer et al. 1996a). The Seitz/GCC group charged that the minor changes made to Chapter 8 during the post-Madrid revision process had somehow dramatically altered the report. Without a shred of evidence, Singer and others asserted that the changes constituted a politically motivated "scientific cleansing."

These irresponsible claims were not reviewed by a single independent, expert peer before being published — in the opinion pages of a business daily and a few news

magazines. We leave it to readers to reflect on how the "flywheel" of peer review might have moderated the assertions of Seitz, Singer, and the GCC.

The Scientific Results Behind the Chapter 8 Conclusions

In a nutshell, the new evidence reported to IPCC and later published in Nature was based not upon new empirical or theoretical results, but on new ways of asking climate models the right questions. In the past, critics such as the University of Virginia's Pat Michaels had correctly argued that direct observational evidence of global warming effects (i.e. "signals") in the climate record were not very well matched to CO2-only model results. For example, CO2-only models suggested that the Earth' surface should have warmed up 1°C rather than the one-half degree C observed in the last century. Additionally, CO2-only models suggested that the Northern Hemisphere would warm up more than the Southern Hemisphere. Such models also, however, suggested the stratosphere would cool as greenhouse gases increased. This clearly was happening, although at least part of that cooling can be attributed to stratospheric ozone depletion (Santer et al. 1996b).

The Earth's warming of a half degree C during the 20th century could be explained simply by asserting the trend to be a natural fluctuation in the climate. The IPCC scientists attempted to estimate the likelihood that natural events were responsible for the observed surface warming. They concluded that this was possible, but improbable. Critics, meanwhile, simply asserted that the warming was natural, without characterizing the probability that this was the correct explanation. Even if it did go unchallenged in a number of op-ed articles, this is a scientifically meaningless claim.

What is the probability that a half-degree warming trend in this century is a natural accident? This cannot be answered by looking at the thermometer record alone, since a globally averaged record is not reliable for much more than a century, if that. It is like trying to determine the probability of "heads" in a coin toss by flipping the coin once. Instead, climate scientists look at proxy records of climate change over long periods of time, such as fluctuating time series of tree ring widths, the deposits left from the comings and goings of glaciers, and the fluctuations of various chemical constituents in ice cores. These records, while not direct measurements of global temperatures, are nonetheless proportional to components of the climate in different parts of the world, and provide a rich record of natural variability.

This record (as summarized in Chapter 3 of the SAR) suggests that the warming of the last century is not unprecedented (Houghton et al. 1996). But it also is not common. Perhaps once in a millennium, such proxy records suggest, a half-degree C global century-long trend could occur naturally (Schneider 1994). In my judgment (SHS), this circumstantial evidence implies that a global surface warming of half a degree has about an 80 to 90 percent likelihood of not being caused by the natural variability of the system. More recent evidence dramatically demonstrates that the last 50 years of the 20th century saw a temperature rise distinctly larger than that of any period in the last 1000 years (Mann et al. 1999).

Natural climatic forcing factors, such as energy output changes on the sun or peculiar patterns of volcanic eruptions, could cause the observed climate trend. However, each

of these climate forcings has a peculiar signature or fingerprint. For example, energy increases from the sun would warm the surface, the lower atmosphere, and the stratosphere all at the same time. On the other hand, greenhouse gas forcing would cool the stratosphere while warming the lower troposphere. Aerosols from human activities, particularly the sulfates generated in coal- and oil-burning regions of the US Northeast, Europe, and China, would cool the troposphere mostly during the day and not at night, and would largely cool the Northern Hemisphere, especially in the summertime when the sun is stronger.

This aerosol effect has turned out to be very important. Indeed, adding sulfate aerosols to greenhouse gas increases in the models led to a dramatic boost in the confidence that could be attached to the circumstantial evidence associated with climatic fingerprints. That is, when the models were driven by both greenhouse gases globally, and sulfate aerosols regionally, no longer did the Northern Hemisphere warm up more than the Southern Hemisphere, or all parts of the high latitudes warm substantially more than the low latitudes. Instead, a different fingerprint pattern emerged. Moreover, this pattern in the models showed an increasing correlation with observations over time — precisely what one would expect in a noisy system in which the human forcing increases with time. By itself, the pattern still has roughly a 10 percent chance of being a random event. However, when taken together with good physical theory and knowledge of ice age-interglacial cycles, seasonal cycles, volcanic eruptions, and now more consistent fingerprints, the vast bulk of the scientific community felt it was not irresponsible to assert that there was a higher likelihood that human climate signals had been detected. Taken together, all this circumstantial evidence was the basis for Chapter 8's now-famous claim that "the balance of evidence suggests a discernible human influence on climate."

At that point in the evolution of knowledge about the Earth's climate system, this was no longer a radical statement. It reflected a lowest-common-denominator consensus view of the vast majority of scientists. It did not say that a climate warming signal had been detected beyond any doubt. Neither we nor any other responsible scientists would make such a claim. But it did offer good reason to begin to plan, responsibly, for the possibility — which we now see as more likely than not — that the global climate will warm by at least one or two degrees during the next 50 years (further support for the likelihood of this outcome appears in Wigley et al. 1998).

The Meaning of "Consensus": Responding to Climate-Change Skeptics

To ignore contrarian critics would be inappropriate. Occasionally, non-conventional outlier opinions revolutionize scientific dogma (Galileo and Einstein are only the most oft-cited examples; see Kuhn 1962). However, we believe that news stories are grossly misleading and irresponsible if they present the unrefereed opinions of skeptics as if they were comparable in credibility to the hundred-scientists, thousand-reviewer documents released by the IPCC. The general public — or lay politicians — cannot be expected to determine for themselves how to weigh these conflicting opinions. And to publish character-assassinating charges of "scientific cleansing" without checking the facts is simply unethical, by the generally accepted standards of scientific practice.

The journalistic doctrine of “balance,” while perhaps appropriate in two-party political systems where the “other side” must always get its equal coverage, is inappropriate if applied literally to multifaceted scientific debates, and it has nothing to do with peer review by experts, especially in the sense we have advocated here. In climate science, wide ranges of probabilities are attached to a whole array of possible outcomes (Morgan and Keith 1995; Nordhaus 1994). Scientific controversy simply cannot be trivialized into a false dichotomy between those who assert that human effects are likely to produce a catastrophic, “end of the world” crisis, “balanced” against those who assert that at worst nothing will happen and at best it will all be good for us. “The end of the world” and “no impact at all” are the two least probable cases (see Schneider 1997).

This is not just a problem for journalists. It also affects scientists. In communication with the public, we sometimes tend to focus our attention on controversies at the cutting edge of the art, rather than present clear perspectives on what is well understood — separating what is truly known from what is merely probable and both of these from what is highly speculative. This, combined with the propensity of the media to focus on “dueling scientists” and extreme, outlier opinions, leads to a miscommunication of the actual nature of the scientific consensus (see the chapter on “Mediarology” in Schneider 1989).

“Consensus,” as we understand it, refers not to a single, exact prediction to which every scientist assents — an impossibility in this field — but to a generally agreed range of possible outcomes. This kind of consensus takes disagreement on details (and even, occasionally, on major points) for granted, both as an unavoidable element of a still-inexact science and as an important motor of scientific progress. Peer review, especially the inclusive and open form adopted by the IPCC, helps to build and maintain it. Consensus of this type is vital to the policy process. In essence, the policy question is to decide how much of current resources should be invested as a hedge against potential negative outcomes. This clearly is a value judgment. It is precisely the kind of judgment that the public and the policy establishment (not scientists) should make, but it can only be made if the decisionmakers — who are not, and are not going to become, experts — are aware of the best range-of-probability and range-of-consequences estimates of the responsible scientific community (see Moss and Schneider 1997).

Faxes sent by special interests to every major journalist on the planet or every significant elected and unelected official — what we like to call the “one fax, one vote” syndrome — are not very good sources of credible knowledge. Vastly better is the work of groups such as the IPCC and the National Research Council, which although slow, deliberative, sometimes elitist, and occasionally dominated by strong personalities, are nonetheless the best representation of the scientific community’s current general opinion.

This kind of scientific consensus is not the same thing as “truth.” Once in a while, the skeptics are right. Indeed, we are certain that some aspects of the current vision of climate change will turn out to be of minor impact, while others will prove to be more serious than currently thought. That is why assessment needs to be a continuous process, and why all policymaking requires “rolling reassessment.” The IPCC, or its progeny, need to be reconvened every five years or so. Only with this input can the political process legitimately decide, and re-decide, to crank up its efforts at mitigation — or to crank them back down, depending upon what is learned in each new assessment about the climate system, the impact of climate change on environment and society, and the effectiveness and distribution of mitigation costs. This ongoing and open process of

refinement of knowledge is the only way that a complex system can become adaptive. Only an adaptive system can minimize the likelihood of making major mistakes, either by overinvesting in environmental protection or by allowing nasty experiments to be performed on “Laboratory Earth” without any attempt to anticipate or slow down the potential negative, irreversible consequences (Schneider 1997).

If the IPCC is to maintain its credibility as a hybrid scientific/political organization, peer review must remain a fundamental formal principle of its self-governance and a basic informal principle of its consensus-building process. Correctly conceived not as a truth machine, but as a technique for improving the quality of science and for moderating the influence of personal, social, and political factors on scientific results, peer review is a powerful technique for generating credible, trusted (and trustworthy) knowledge. The IPCC’s widely inclusive, extremely intensive peer review process has opened the debate about climate change to a far wider range of actors than is usually consulted in science. By doing so, it has created a fairer, more thorough, and hence more powerful method for reaching consensus on the knowledge required for good public policy.

References

Abelson, P. H. (1980) "Scientific Communication," Science 209 (4452): 60-62.

Bolin, B. (1996) Letter to Ben Santer: Original letter quoted here (provided by Bert Bolin). A brief excerpt from the full letter was published by the Wall Street Journal on 25 June.

Bolin, B., J. Houghton, and L. G. Meira-Filho (1996) "Letter to the Editor," Wall Street Journal, June 25.

Bornstein, R. F. (1991) "The Predictive Validity of Peer Review: A Neglected Issue," Behavioral and Brain Science 14 (1): 138-139.

Brown, G. E., Rep. (1996) "Environmental Science Under Siege: Fringe Science and the 104th Congress." Report, Democratic Caucus of the Committee on Science, U.S. House of Representatives, October 23, Washington, D.C. Available at http://www.house.gov/science_democrats/archive/envrpt96.htm.

Bruce, J., H. Lee, and E. Haites, eds. (1996) Climate Change 1995: Economic and Social Dimensions of Climate Change, Vol. 3 of 3 (Cambridge, UK: Cambridge University Press).

Chubin, D. E. and E. J. Hackett (1990) Peerless Science: Peer Review and U.S. Science Policy, SUNY series in science, technology, and society (Albany, N.Y.: State University of New York Press).

Cichetti, D. V. (1991) "The Reliability of Peer Review for Manuscript and Grant Submissions: A Cross-Disciplinary Investigation," Behavioral and Brain Sciences 14: 119-135.

Cole, S. (1992) Making Science: Between Nature and Society (Cambridge, MA: Harvard University Press).

Collins, H. and T. Pinch (1993) The Golem: What Everyone Should Know about Science (Cambridge: Cambridge University Press).

Daniel, H.-D. (1993) Guardians of Science: Fairness and Reliability of Peer Review, translated by Russey, William E. (New York: VCH).

Foucault, M. (1977) Discipline and Punish, translated by Alan Sheridan (New York: Vintage Books).

Gilbert, G. N. and M. Mulkay (1984) Opening Pandora's Box: A Sociological Analysis of Scientists' Discourse (Cambridge: Cambridge University Press).

Haas, P. M. (1990) "Obtaining International Environmental Protection through Epistemic Consensus," Millennium 19 (3): 347-364.

Houghton, J. J. et al., eds. (1996) Climate Change 1995: The Science of Climate Change, Vol. 1 of 3 (Cambridge, UK: Cambridge University Press).

Intergovernmental Panel on Climate Change (1993) "IPCC Procedures for Preparation, Review, Acceptance, Approval, and Publication of its Reports." Produced by IPCC Secretariat. Last updated June 30, 1993. Available at http://www.usgcrp.gov/ipcc/html/rulespro.html.

Jasanoff, S. (1990) The Fifth Branch: Science Advisors as Policymakers (Cambridge, MA: Harvard University Press).

Jasanoff, S. (1991) "Acceptable Evidence in a Pluralistic Society," in D. G. Mayo and R. D. Hollander, eds., Acceptable Evidence: Science and Values in Risk Management (New York: Oxford University Press).

Jasanoff, S. et al. (1998) "Science and Decisionmaking," in S. Rayner and E. L. Malone, eds., Human Choice and Climate Change: The Societal Framework (Columbus, Ohio: Battelle Press).

Judson, H. F. (1994) "Structural transformation of the sciences and the end of peer review," Journal of the American Medical Association 272: 92-94.

Kassirer, J. P. and E. W. Campion (1994) "Peer review: crude and understudied, but indispensable," Journal of the American Medical Association 272: 96-97.

Kuhn, T. S. (1962) The Structure of Scientific Revolutions (Chicago: University of Chicago).

Latour, B. and S. Woolgar (1979) Laboratory Life: The Social Construction of Scientific Facts (London: Sage).

Leary, N. (1999) "Non-peer reviewed sources." Electronic mail to IPCC authors. Technical Support Unit, IPCC Working Group II, July 19.

Mahoney, M. J. (1977) "Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System," Cognitive Therapy and Research 1: 161-175.

Mann, M. E., R. S. Bradley, and M. K. Hughes (1999) "Northern Hemisphere Temperatures During the Past Millennium: Inferences, Uncertainties, and Limitations," Geophysical Research Letters 29 (6): 759.

Marsh, H. W. and S. Ball (1989) "The Peer Review Process Used to Evaluate Manuscripts Submitted to Academic Journals: Interjudgmental Reliability," Journal of Experimental Education 57 (2): 151-169.

Masood, E. (1996) "Climate Report 'Subject to Scientific Cleansing'," Nature 381 (6583): 546.

Merton, R. K. (1973) The Sociology of Science: Theoretical and Empirical Investigations (Chicago: University of Chicago Press).

Moran, G. (1998) Silencing Scientists and Scholars in Other Fields: Power, Paradigm Controls, Peer Review, and Scholarly Communication, Contemporary studies in information management, policy, and services (Greenwich, Ct.: Ablex Publishing).

Morgan, M. G. and D. W. Keith (1995) "Subjective Judgments by Climate Experts," Environmental Science and Technology 29: 468A-476A.

Moss, R. and S. H. Schneider (1997) "Session Synthesis Essay: Characterizing and Communicating Scientific Uncertainty: Building on the IPCC Second Assessmen," in S. J. Hassol and J. Katzenberger, eds., Elements of Change 1996 (Aspen, CO: Aspen Global Change Institute). Available at http://www.gcrio.org/ASPEN/science/eoc96/AGCIEOC96SSSII/AGCIEOC96SynthesisSSSII.html.

Nature editors (1996) "Climate Debate Must Not Overheat," Nature 381 (6583): 539.

Nordhaus, W. D. (1994) "Expert Opinion on Climatic Change," American Scientist 82 (1): 45-51.

Ross, P. F. (1980) The Sciences' Self-Management: Manuscript Refereeing, Peer Review, and Goals in Science (Lincoln, MA: The Ross Company).

Roy, R. (1985) "Funding Science: The Real Defects of Peer Review and an Alternative to It," Science, Technology, & Human Values 10 (3).

Santer, B. D. et al. (1996b) "A Search for Human Influences on the Thermal Structure of the Atmosphere," Nature 382 (6586): 39-46.

Santer, B. et al. (1996a) "Response to Wall Street Journal Editorial of June 12th, 1996 by Frederick Seitz," Wall Street Journal, June 25.

Schneider, S. H. (1989) Global Warming: Are We Entering the Greenhouse Century? (New York: Vintage Books).

Schneider, S. H. (1994) "Detecting Climatic Change Signals: Are There Any 'Fingerprints'?," Science, January 21: 341-347.

Schneider, S. H. (1997) Laboratory Earth: The Planetary Gamble We Can't Afford To Lose (New York: Basic Books).

Seitz, F. (1996) "A Major Deception on Global Warming," Wall Street Journal, June 12.

Shapin, S. and S. Shaffer (1985) Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life (Princeton: Princeton University Press).

Singer, S. F. (1996) "Letter to the Editor," Wall Street Journal, July 11.

Wamsted, D. (1996) "Doctoring the Documents?," Energy Daily 24 (98): 1-2.

Watson, R. T., M. C. Zinyowera, and R. H. Moss, eds. (1996) Climate Change 1995: Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses, Vol. 2 of 3 (Cambridge, UK: Cambridge University Press).

Wigley, T. M. L., R. L. Smith, and B. D. Santer (1998) "Anthropogenic influence on the autocorrelation function of hemispheric-mean temperatures," Science 282: 1676-1679.

Wirth, T. E. (1996) "Statement on behalf of the United States of America." Provided by USGCRP office. Second Conference of Parties to the Framework Convention on Climate Change, July 17, Geneva, Switzerland.

Notes

[i] A previous version of this article was published as P. N. Edwards and S. H. Schneider, "The IPCC 1995 Report: Broad Consensus or 'Scientific Cleansing'?," *Ecofables/Ecoscience* 1:1 (1997), 3-9. The authors wish to thank Simon Shackley, Ben Santer, and Michael Oppenheimer for their helpful comments on an early draft.

[ii] The new rules, once finalized, will be available at http://www.ipcc.ch/.

[iii] "Virtual witnesses" "watch" a scientific experiment by reading its written description, "witnessing" it second-hand and validating it by agreeing that it was correctly performed and that the reasoning used to reach conclusions was correct.

[iv] See Miller, Chapter 8, this volume, for a more negative view of the IPCC's inclusiveness.