CLIMATE SCIENCE AND THE UNCERTAINTY MONSTER

BY J. A. CURRY AND P. J. WEBSTER

An exploration of ways to understand, assess and reason about uncertainty in climate science, with specific application to the IPCC assessment process.

Doubt is not a pleasant condition, but certainty is absurd.

—VOLTAIRE

Over the course of history, what seems unknowable and unimaginable to one generation becomes merely a technical challenge for a subsequent generation. The "endless frontier" of science (Bush 1945) advances as scientists extend what is possible both in theory and practice. Doubt and uncertainty about our current understanding is inherent at the knowledge frontier. While extending the knowledge frontier often reduces uncertainty, it leads inevitably to greater uncertainty as unanticipated complexities are discovered. A scientist's perspective of the knowledge frontier is described by Feynman (1988): "When a scientist does not know the answer to a problem, he is ignorant. When he has a hunch as to what the result is, he is uncertain. And when he is pretty damn sure of what the result is going to be, he is still in some doubt. We have found it of paramount importance that in order to progress, we must recognize our ignorance and leave room for doubt. Scientific knowledge is a body of statements of varying degrees of certainty—some most unsure, some nearly sure, but none absolutely certain."

How to understand and reason about uncertainty in climate science is a topic that is receiving increasing attention in both the scientific and philosophical literature. Such inquiry is paramount because of the challenges to climate science associated with the science–policy interface and its socioeconomic importance, as reflected by the Intergovernmental Panel for Climate Change (IPCC) assessment reports (all IPCC assessment reports are available online at www.ipcc.ch/publications_and_data/publications_and_data_reports.htm#1.)[1]

The "uncertainty monster" is a concept introduced by Van der Sluijs (2005) in an analysis of the different ways that the scientific community responds to uncertainties that are difficult to tame. The "monster" is the

AFFILIATIONS: CURRY AND WEBSTER—School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia
CORRESPONDING AUTHOR: Judith Curry, School of Earth and Atmospheric Sciences, Georgia Institute of Technology, 311 Ferst Drive, Atlanta, GA 30332
E-mail: curryja@eas.gatech.edu

The abstract for this article can be found in this issue, following the table of contents.
DOI:10.1175/2011BAMS3139.1

In final form 20 June 2011

[1] The first–fourth assessment reports (ARs) are referred to here as FAR, SAR, TAR, AR4, plus the forthcoming AR5. Unless otherwise indicated, citations in the text refer to Working Group I reports.

confusion and ambiguity associated with knowledge versus ignorance, objectivity versus subjectivity, facts versus values, prediction versus speculation, and science versus policy. The uncertainty monster gives rise to discomfort and fear, particularly with regard to our reactions to things or situations we cannot understand or control, including the presentiment of radical unknown dangers. An adaptation of Van der Sluijs's strategies of coping with the uncertainty monster at the science–policy interface is described below.

- *Monster hiding.* Uncertainty hiding or the "never admit error" strategy can be motivated by a political agenda or because of fear that uncertain science will be judged as poor science by the outside world. Apart from the ethical issues of monster hiding, the monster may be too big to hide and uncertainty hiding enrages the monster.
- *Monster exorcism.* The uncertainty monster exorcist focuses on reducing the uncertainty through advocating for more research. In the 1990s, a growing sense of the infeasibility of reducing uncertainties in global climate modeling emerged in response to the continued emergence of unforeseen complexities and sources of uncertainties. Van der Sluijs (2005, p. 88) states that "monster-theory predicts that [reducing uncertainty] will prove to be vain in the long run: for each head of the uncertainty monster that science chops off, several new monster heads tend to pop up due to unforeseen complexities," analogous to the Hydra beast of Greek mythology.
- *Monster simplification.* Monster simplifiers attempt to transform the monster by subjectively quantifying and simplifying the assessment of uncertainty. Monster simplification is formalized in the IPCC TAR and AR4 by guidelines for characterizing uncertainty in a consensus approach consisting of expert judgment in the context of a subjective Bayesian analysis (Moss and Schneider 2000).
- *Monster detection.* The first type of uncertainty detective is the scientist who challenges existing theses and works to extend knowledge frontiers. The second type is the watchdog auditor, whose main concern is accountability, quality control, and transparency of the science. The third type is the merchant of doubt (Oreskes and Collins 2010), who distorts and magnifies uncertainties as an excuse for inaction for financial or ideological reasons.
- *Monster assimilation.* Monster assimilation is about learning to live with the monster and giving uncertainty an explicit place in the contemplation and management of environmental risks. Assessment and communication of uncertainty and ignorance, along with extended peer communities, are essential in monster assimilation. The challenge to monster assimilation is the ever-changing nature of the monster and the birth of new monsters.

This paper explores ways to understand, assess, and reason about uncertainty in climate science, with specific application to the IPCC assessment process. Section 2 describes the challenges of understanding and characterizing uncertainty in dynamical models of complex systems, including challenges to interpreting the ensemble of simulations for the twenty-first-century climate used in the IPCC assessment reports. Section 3 addresses some issues regarding reasoning about uncertainty and examines the treatment of uncertainty by the IPCC Assessment Reports. Section 4 addresses uncertainty in the detection and attribution of anthropogenic climate change. And finally, section 5 introduces some ideas for monster taming strategies at the levels of institutions, individual scientists, and communities.

UNCERTAINTY OF CLIMATE MODELS.

Synergy means behavior of whole systems unpredicted by the behavior of their parts.

—R. Buckminster Fuller

Climate model complexity arises from the nonlinearity of the equations' high dimensionality (millions of degrees of freedom) and the linking of multiple subsystems. Computer simulations of the complex climate system can be used to represent aspects of climate that are extremely difficult to observe, experiment with theories in a new way by enabling hitherto infeasible calculations, understand a system of equations that would otherwise be impenetrable, and explore the system to identify unexpected outcomes (e.g., Muller 2010).

Imperfect models.

The future ain't what it used to be.

—Yogi Berra

Model imperfection is a general term that describes our limited ability to simulate climate and is categorized here in terms of model inadequacy and model uncertainty. Model inadequacy reflects our limited understanding of the climate system, inadequacies of numerical solutions employed in computer models, and the fact that no model can be structurally identical to the actual system (e.g., Stainforth et al. 2007). Model structural form is the conceptual modeling of the physical system (e.g., dynamical equations, initial and boundary conditions), including the selection of subsystems to include (e.g., stratospheric chemistry, ice sheet dynamics). In addition to insufficient understanding of the system, uncertainties in model structural form are introduced as a pragmatic compromise between numerical stability and fidelity to the underlying theories, credibility of results, and available computational resources.

UNCERTAINTY LEXICON

The nature of uncertainty is often expressed by the distinction between epistemic uncertainty and ontic uncertainty.

Epistemic uncertainty is associated with imperfections of knowledge, which may be reduced by further research and empirical investigation. Examples include limitations of measurement devices and insufficient data. Epistemic uncertainties in models include missing or inadequately treated processes and errors in the specification of boundary conditions.

Ontic (often referred to as *aleatory*) *uncertainty* is associated with inherent variability or randomness.

Natural internal variability of the climate system contributes to ontic uncertainty in the climate system. Ontic uncertainties are by definition irreducible.

Walker et al. (2003) provides a complete logical structure of the level of uncertainty, characterized as a progression between deterministic understanding and total ignorance: statistical uncertainty, scenario uncertainty, and recognized ignorance.

Statistical uncertainty is the aspect of uncertainty that is described in statistical terms. An example of statistical uncertainty is measurement uncertainty, which can be due to sampling error or inaccuracy or imprecision in measurements.

Scenario uncertainty implies that it is not possible to formulate the probability of occurrence of one particular outcome. A scenario is a plausible but unverifiable description of how the system and/or its driving forces may develop over time. Scenarios may be regarded as a range of discrete possibilities with no a *priori allocation* of likelihood.

Recognized ignorance refers to fundamental uncertainty in the mechanisms being studied and a weak scientific basis for developing scenarios. *Reducible ignorance* may be resolved by conducting further research, whereas *irreducible ignorance* implies that research cannot improve knowledge.

An alternative taxonomy for levels of uncertainty is illustrated by this quote from U.S. Secretary of Defense Donald Rumsfeld (U.S. DOD 2011): "[A]s we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns—the ones we do not know we do not know. And if one looks throughout the history of our country and other free countries, it is the latter category that tend to be the difficult ones."

Model uncertainty is associated with uncertainty in model parameters and subgrid parameterizations, and also with uncertainty in initial conditions. Uncertainties in parameter values include uncertain constants and other parameters that are largely contained in subgrid-scale parameterizations (e.g., boundary layer turbulence, cloud microphysics), and parameters involved in ad hoc modeling to compensate for the absence of neglected factors. Initial condition uncertainty arises in simulations of nonlinear and chaotic dynamical systems (e.g., Palmer et al. 2005). If the initial conditions are not known exactly, then the forecast trajectory will diverge from the actual trajectory, and it cannot be assumed that small perturbations have small effects. As such, model uncertainty includes epistemic uncertainty in parameter values and both epistemic and ontic uncertainty in initial conditions.

Ensemble methods are a brute force approach to representing model parameter and initial condition uncertainty (for an overview, see Parker 2010). Rather than conducting a single simulation, multiple simulations are run that sample some combination of different initial conditions, model parameters and parameterizations, and model structural forms. While the ensemble method used in weather and climate predictions is inspired by Monte Carlo approaches, the application of a traditional Monte Carlo approach far outstrips computational capacity owing to the very large number of possible combinations required to fully represent climate model parameter and initial condition uncertainty. A high level of model complexity and high model resolution precludes large ensembles. Stochastic parameterization methods are being introduced (e.g., Palmer 2001) to characterize parameter and parameterization uncertainty, reducing the need to conduct ensemble simulations to explore parameter and parameterization uncertainty.

Model outcome uncertainty, also referred to as prediction error, arises from the propagation of the aforementioned uncertainties through the model simulation and is evidenced by the simulated outcomes. Model prediction error can be evaluated against known analytical solutions, comparisons with other simulations, and/or comparison with observations. Reducing prediction error is a fundamental objective of model calibration. Calibration is necessary to address parameters that are unknown or inapplicable at the model resolution, and also in the linking of submodels. As the complexity, dimensionality, and modularity of a model grow, model calibration becomes unavoidable and an increasingly important issue. Model calibration is accomplished by kludging (or tuning), which is "an inelegant, botched together piece of program; something functional but somehow messy and unsatisfying, a piece of program or machinery which works up to a point" (Lenhard and Winsberg 2011, p. 121). A kludge required in one model may not be required in another model that has greater structural adequacy or higher resolution. Continual ad hoc adjustment of the model (calibration) provides a means for the model to avoid being falsified; Occam's razor presupposes that the model least dependent on continual ad hoc modification is to be preferred.

A serious challenge to improving complex nonlinear models is that model complexity and analytic impenetrability precludes the precise evaluation of the location of parameter(s) that are producing the prediction error (Lenhard and Winsberg 2010). For example, if a model is producing shortwave surface radiation fluxes that are substantially biased relative to observations, it is impossible to determine whether the error arises from the radiative transfer model, incoming solar radiation at the top of the atmosphere, concentrations of the gases that absorb shortwave radiation, physical and chemical properties of the aerosols in the model, morphological and microphysical properties of the clouds, convective parameterization that influences the distribution of water vapor and clouds, and/or characterization of surface reflectivity. Whether a new parameterization module adds to or subtracts from the overall reliability of the model may have more to do with some entrenched features of model calibration than it does with that module's fidelity to reality when considered in isolation.

Confidence and credibility.

All models are wrong, but some are useful.

—George E. P. Box

Confidence is a degree of certainty that a particular model is effective or useful. Confidence is inspired by the model's relation to theory and physical understanding of the processes involved, sensitivity of the simulations to model structure, the nature of the ad hoc adjustments and calibration, extensive exploration of model uncertainty, consistency of the simulated responses, and the ability of the model and model components to simulate historical observations (e.g., Knutti 2008). User confidence in a forecast model system depends critically on the confirmation of forecasts, both using historical data (hindcasts, in-sample) and actual forecasts (out-of-sample observations). Parker (2009) argues that instances of fit between model output and observational data do not confirm the models themselves, but rather hypotheses about the adequacy of climate models for particular purposes. Hence, model validation strategies depend on the intended application of the model. However, there is no generally agreed upon protocol for the validation of climate models (e.g., Guillemot 2010).

User confidence in a forecast model depends critically on the confirmation of forecasts, both using historical data (hindcasts, in-sample) and out-of-sample observations (forecasts). Confirmation with out-of-sample observations is possible for forecasts that have a short time horizon that can be compared with out-of-sample observations (e.g., weather forecasts). Unless the model can capture or bound a phenomenon in hindcasts and previous forecasts, there is no expectation that the model can quantify the same phenomena in subsequent forecasts. Capturing the phenomena in hindcasts and previous forecasts does not in any way guarantee the ability of the model to capture the phenomena in the future, but it is a necessary condition (Smith 2002). If the distance of future simulations from the established range of model validity is small, then it reasonable to extend established confidence in the model to the perturbed future state. Extending such confidence requires that no crucial feedback mechanisms are missing from the model (Smith 2002).

Even for in-sample validation, there is no straightforward definition of model performance for complex nondeterministic models having millions of degrees of freedom (e.g., Guillemot 2010). Because the models are not deterministic, multiple simulations are needed to compare with observations, and the number of simulations conducted by modeling centers are insufficient to establish a robust mean; hence, bounding box approaches (assessing whether the range of the ensembles bounds the observations; Judd et al. 2007) are arguably a better way to establish empirical adequacy. A further complication arises if datasets used in the model evaluation process are the same as those used for calibration, which gives rise to circular reasoning (confirming the antecedent) in the evaluation process.

On the subject of confidence in climate models, Knutti (2008, p. 2654) summarizes, "So the best we can hope for is to demonstrate that the model does not violate our theoretical understanding of the system and that it is consistent with the available data within the observational uncertainty."

Simulations of the twenty-first-century climate.

There are many more ways to be wrong in a 10^6 dimensional space than there are ways to be right.

—Leonard Smith

What kind of confidence can we have in the simulations of scenarios for the twenty-first century? Since projections of future climate relate to a state of the system that is outside the range of model validity, it is therefore impossible to either calibrate the model for the forecast regime of interest or confirm the usefulness of the forecasting process. The problem is further exacerbated by the lifetime of an individual model version being substantially less than the prediction lead time (Smith 2002).

If the distance of future simulations from the established range of model validity is small, then it reasonable to extend established confidence in the model to the perturbed future state. In effect, such confidence requires that we assume that nothing happens that takes the model farther beyond its range of validity, and that no crucial feedback mechanisms are missing from the model (Smith 2002). Of particular relevance to simulations with increased greenhouse gases is the possibility that slow changes in the forcing may push the model beyond a threshold and induce a transition to a second equilibrium.

A key issue in assessing model adequacy for twenty-first-century climate simulations is the inclusion of longer time-scale processes, such as the global carbon cycle and ice sheet dynamics. In addition to these known unknowns, there are other processes that we have some hints of but currently have no way of quantifying (e.g., methane release from thawing permafrost). Confidence established in the atmospheric dynamical core as a result of the extensive cycles of evaluation and improvement of weather forecast models is important, but other factors become significant in climate models that have less import in weather models, such as mass conservation and cloud and water vapor feedback processes.

Given the inadequacies of current climate models, how should we interpret the multimodel ensemble simulations of the twenty-first-century climate used in the IPCC assessment reports? This ensemble of opportunity is composed of models with generally similar structures but different parameter choices and calibration histories (for an overview, see Knutti et al. 2008; Hargreaves 2010). McWilliams (2007) and Parker (2010) argue that current climate model ensembles are not designed to sample representational uncertainty in a thorough or strategic way. Stainforth et al. (2007) argue that model inadequacy and an inadequate number of simulations in the ensemble preclude producing meaningful probability density functions (PDFs) from the frequency of model outcomes of future climate. Nevertheless, as summarized by Parker (2010), it is becoming increasingly common for results from individual multimodel and perturbed physics simulations to be transformed into probabilistic projections of future climate, using Bayesian and other techniques. Parker argues that the reliability of these probabilistic projections is unknown, and in many cases they lack robustness. Knutti et al. (2008) argues that the real challenge lies more in how to interpret the PDFs than in whether they should be constructed in the first place. Stainforth et al. (2007) warns against overinterpreting current model results since they could be contradicted by the next generation of models, undermining the credibility of the new generation of model simulations.

Stainforth et al. (2007) emphasize that models can provide useful insights without being able to provide probabilities, by providing a lower bound on the maximum range of uncertainty and a range of possibilities to be considered. Kandlikar et al. (2005) argue that when sources of uncertainty are well understood, it can be appropriate to convey uncertainty via full PDFs; however, in other cases, it will be more appropriate to offer only a range in which one expects the value of a predictive variable to fall with some specified probability, or to indicate the expected sign of a change without assigning a magnitude. They argue that uncertainty should be expressed using the most precise means that can be justified, but that unjustified more precise means should not be used.

UNCERTAINTY AND THE IPCC.

You are so convinced that you believe only what you believe that you believe, that you remain utterly blind to what you really believe without believing you believe it.

—Orson Scott Card, *Shadow of the Hegemon*

How to reason about uncertainties in the complex climate system and its computer simulations is not simple or obvious. Scientific debates involve controversies over the value and importance of particular classes of evidence as well as disagreement about the appropriate logical framework for linking and assessing the evidence. The IPCC faces a daunting challenge with regard to characterizing and reasoning about uncertainty, assessing the quality of evidence, linking the evidence into arguments, identifying areas of ignorance, and assessing confidence levels.

Characterizing uncertainty.

A long time ago a bunch of people reached a general consensus as to what's real and what's not and most of us have been going along with it ever since.

—Charles de Lint

Over the course of four assessment reports, the IPCC has given increasing attention to reporting uncertainties (e.g., Swart et al. 2009). The "guidance paper" by Moss and Schneider (2000) recommended steps for assessing uncertainty in the IPCC assessment reports and a common vocabulary to express quantitative levels of confidence based on the amount of evidence (number of sources of information) and the degree of agreement (consensus) among experts (see sidebar for vocabulary).

The IPCC guidance for characterizing uncertainty for the AR4 (WMO 2005) describes three approaches for indicating confidence in a particular result and/or that the likelihood that a particular conclusion is correct:

1) A qualitative level-of-understanding scale describes the level of scientific understanding in terms of the amount of evidence available and the degree of agreement among experts. There can be limited, medium, or much evidence, and agreement can be low, medium, or high.
2) A quantitative confidence scale estimates the level of confidence for a scientific finding and ranges from "very high confidence" (9 in 10 chance) to "very low confidence" (less than 1 in 10 chance).
3) A quantitative likelihood scale represents "a probabilistic assessment of some well-defined outcome having occurred or occurring in the future." The scale ranges from "virtually certain" (greater than 99% probability) to "exceptionally unlikely" (less than 1% probability).

Oppenheimer et al. (2007), Webster (2009), Petersen (2006), and Kandlikar et al. (2005) argue that future IPCC efforts need to be more thorough about describing sources and types of uncertainty, making the uncertainty analysis as transparent as possible. The InterAcademy Council (IAC; http://reviewipcc.interacademycouncil.net/) reviewed the IPCC's performance on characterizing uncertainty. In response to concerns raised in the review, the IAC made the following recommendations regarding the IPCC's treatment of uncertainty:

- "Each Working Group should use the qualitative level-of-understanding scale in its Summary for Policymakers and Technical Summary, as suggested in IPCC's uncertainty guidance for the Fourth Assessment." This is a key element of uncertainty monster detection.
- "Chapter Lead Authors should provide a traceable account of how they arrived at their ratings for level of scientific understanding and likelihood that an outcome will occur." Failure to provide a traceable account is characteristic of uncertainty monster hiding.
- "Quantitative probabilities (as in the likelihood scale) should be used to describe the probability of well-defined outcomes only when there is sufficient evidence. Authors should indicate the basis for assigning a probability to an outcome or event (e.g., based on measurement, expert judgment, and/or model runs)." Using quantitative probabilities when there is insufficient evidence is uncertainty monster simplification.

The recommendations made by the IAC concerning the IPCC's characterization of uncertainty are steps in the right direction in terms of dealing with the uncertainty monster. Curry (2011a) further argued that a concerted effort by the IPCC is needed to identify better ways of framing the climate change problem, exploring and characterizing uncertainty, reasoning about uncertainty in the context of evidence-based logical hierarchies, and eliminating bias from the consensus building process itself.

Reasoning about uncertainty.

It is not so much that people hate uncertainty, but rather that they hate losing.

—Amos Tversky

The IPCC characterization of characterization is based upon a consensus building process that is an exercise in collective judgment in areas of uncertain knowledge. The general reasoning underlying the IPCC's arguments for anthropogenic climate change combines a compilation of evidence with subjective Bayesian reasoning. This process is described by Oreskes (2007) as presenting a "consilience of evidence" argument, which consists of independent lines of evidence that are explained by the same theoretical account.

Given the complexity of the climate problem, expert judgments about uncertainty and confidence levels are made by the IPCC on issues that are dominated by unquantifiable uncertainties. Curry (2011a) argues that because of the complexity of the issues, individual experts use different mental models for evaluating the interconnected evidence. Biases can abound when reasoning and making judgments about such a complex problem. Bias can occur as a result of excessive reliance on a particular piece of evidence, the presence of cognitive biases in heuristics, failure to account for indeterminacy and ignorance, and logical fallacies and errors, including circular reasoning. The IAC (2010, p. 41) states that "studies suggest that informal elicitation measures, especially those designed to reach consensus, lead to different assessments of probabilities than formal measures. Informal procedures often result in probability distributions that place less weight in the tails of the distribution than formal elicitation methods, possibly understating the uncertainty associated with a given outcome."

Oreskes (2007) draws an analogy for the consilience of evidence approach with what happens in a legal case. Continuing with the legal analogy, Johnston (2010) characterized the IPCC's arguments as a legal brief, designed to persuade, in contrast to a legal memo that is intended to objectively assess both sides. Along the lines of a legal memo, Curry (2011a) argues that the consilience of evidence argument is not convincing unless it includes parallel evidence-based analyses for competing hypotheses, and hence a critical element in uncertainty monster detection. Any evidence-based argument that is more inclined to admit one type of evidence or argument rather than another tends to be biased. Parallel evidence-based analysis of competing hypotheses provides a framework whereby scientists with a plurality of viewpoints participate in an assessment. In a Bayesian analysis with multiple lines of evidence, it is conceivable that there are multiple lines of evidence that produce a high confidence level for each of two opposing arguments, which is referred to as the ambiguity of competing certainties. If uncertainty and ignorance are acknowledged adequately, then the competing certainties disappear. Disagreement then becomes the basis for focusing research in a certain area, and so moves the science forward.

UNCERTAINTY IN THE ATTRIBUTION OF TWENTIETH-CENTURY CLIMATE CHANGE.

Give me four parameters, and I can fit an elephant. Give me five, and I can wiggle its trunk.

—John von Neumann

Arguably the most important conclusion of IPCC AR4 is the following statement: "Most of the observed increase in global average temperatures since the mid-20th century is *very likely* due to the observed increase in anthropogenic greenhouse gas concentrations" (IPCC 2007, p. 10). This section raises issues regarding the uncertainties that enter into the attribution argument, ambiguities in the attribution statement and apparent circular reasoning, and lack of traceability of the "very likely" likelihood assessment.

IPCC's detection and attribution argument.

What we observe is not nature itself, but nature exposed to our method of questioning.

—Werner Karl Heisenberg

The problem of attributing climate change is intimately connected with the detection of climate change. A change in the climate is "detected" if its likelihood of occurrence by chance due to internal variability alone is determined to be small. Knowledge of internal climate variability is needed for both detection and attribution. Because the instrumental record is too short to give a well-constrained estimate of internal variability, internal climate variability is usually estimated from long control simulations from coupled climate models. The IPCC AR4 (Hegerl et al. 2007, p. 668) formulates the problem of attribution to be: "In practice attribution of anthropogenic climate change is understood to mean demonstration that a detected change is 'consistent with the estimated responses to the given combination of anthropogenic and natural forcing' and 'not consistent with alternative, physically plausible explanations of recent climate change that exclude important elements of the given combination of forcings'" (Mitchell et al. 2001, p. 700).

Detection and attribution analyses use objective statistical tests to assess whether observations contain evidence of the expected responses to external forcing that is distinct natural internal variability. Expected responses, or "fingerprints," are determined from climate models and physical understanding of the climate system. Formal Bayesian reasoning is used to some extent by the IPCC in making inferences about detection and attribution. The reasoning process used in assessing likelihood in the attribution statement is described by this statement from the AR4 (Hegerl et al. 2007, p. 669):

> The approaches used in detection and attribution research described above cannot fully account for all uncertainties, and thus ultimately expert judgment is required to give a calibrated assessment of whether a specific cause is responsible for a given

climate change. The assessment approach used in this chapter is to consider results from multiple studies using a variety of observational data sets, models, forcings and analysis techniques. The assessment based on these results typically takes into account the number of studies, the extent to which there is consensus among studies on the significance of detection results, the extent to which there is consensus on the consistency between the observed change and the change expected from forcing, the degree of consistency with other types of evidence, the extent to which known uncertainties are accounted for in and between studies, and whether there might be other physically plausible explanations for the given climate change. Having determined a particular likelihood assessment, this was then further downweighted to take into account any remaining uncertainties, such as, for example, structural uncertainties or a limited exploration of possible forcing histories of uncertain forcings. The overall assessment also considers whether several independent lines of evidence strengthen a result.

The IPCC AR4 (Hegerl et al. 2007) describes two types of simulation methods that have been used in detection and attribution studies. The first method is a "forward calculation" that uses best estimates of external changes in the climate system (forcings) to simulate the response of the climate system using a climate model. These forward calculations are then directly compared to the observed changes in the climate system. The second method is an "inverse calculation," whereby the magnitude of uncertain model parameters and applied forcing is varied to provide a best fit to the observational record. While the exact reasoning underlying the IPCC's likelihood assessment is unclear, the important role of coupled climate models in the assessment is indicated by the fact that 12 of the 14 figures in sections 9.2–9.4 in Hegerl et al. (2007) are based upon the results of climate model simulations.

Whereas all of the climate model simulations and various attribution studies agree that the warming observed since 1970 can only be reproduced using anthropogenic forcings, models and attribution analyses disagree on the relative importance of solar, volcanic, and aerosol forcing in the earlier part of the twentieth century (section 9.4.1 in Hegerl et al. 2007). The substantial warming during the period 1910–40 has been attributed by nearly all the modeling groups to some combination of increasing solar irradiance and a lack of major volcanic activity. The cooling and leveling off of average global temperatures during the 1950s and 1960s is attributed primarily to aerosols from fossil fuels and other sources, when the greenhouse warming was overwhelmed by aerosol cooling.

Sources of uncertainty.

Not only does God play dice, but sometimes he throws the dice where we can't see them.

—Stephen Hawking

Attribution of observed climate change is affected by errors and uncertainties in the prescribed external forcing and in the model's capability to simulate both the response to the forcing (sensitivity) and decadal-scale natural internal variability. Uncertainties in the model and forcing are acknowledged by the AR4 (Hegerl et al. 2007, p. 669): "Ideally, the assessment of model uncertainty should include uncertainties in model parameters (e.g., as explored by multi-model ensembles), and in the representation of physical processes in models (structural uncertainty). Such a complete assessment is not yet available, although model intercomparison studies (chapter 8) improve the understanding of these uncertainties. The effects of forcing uncertainties, which can be considerable for some forcing agents such as solar and aerosol forcing (section 9.2), also remain difficult to evaluate despite advances in research."

The level of scientific understanding of radiative forcing is ranked by the AR4 (Table 2.11 in Forster et al. 2007) as high only for the long-lived greenhouse gases, but it is ranked as low for solar irradiance, aerosol effects, stratospheric water vapor from CH_4, and jet contrails. Radiative forcing time series for the natural forcings (solar, volcanic aerosol) are reasonably well known for the past 25 years, with estimates farther back in time having increasingly large uncertainties.

Based upon new and more reliable solar reconstructions, the AR4 (Forster et al. 2007, section 2.7.1.2) concluded that the increase in solar forcing during the period 1900–80 used in the AR3 reconstructions is questionable and that the direct radiative forcing due to an increase in solar irradiance is reduced substantially by the AR3. However, consideration of Table S9.1 in the Hegerl et al. (2007) shows that each climate model used outdated solar forcing (from the AR3) that was assessed to substantially overestimate the magnitude of the trend in solar forcing prior to 1980. The IPCC AR4 (Hegerl et al. 2007, p. 679) states that "while the 11-year solar forcing cycle is well documented, lower-frequency variations in solar forcing are highly uncertain." Furthermore, "large uncertainties associated with estimates of past solar forcing (section 2.7.1) and omission of some chemical and dynamical response mechanisms (Gray et al., 2005) make it difficult to reliably estimate the contribution of solar forcing to warming over the 20th century."

The greatest uncertainty in radiative forcing is associated with aerosols, particularly the aerosol indirect effect, whereby aerosols influence cloud radiative properties. Consideration of Fig. 2.20 of the AR4 (Forster et al. 2007) shows that, given the uncertainty in aerosol forcing, the magnitude of the aerosol forcing (which is negative, or cooling) could rival the forcing from long-lived greenhouse gases (positive, or warming). The twentieth-century aerosol forcing used in most of the AR4 model simulations (Forster et al. 2007, section 9.2.1.2) relies on inverse calculations of aerosol optical properties to match climate model simulations with observations. The only constraint on the aerosol forcing used in the AR4 attribution studies is that the derived forcing should be within the bounds of forward calculations that determine aerosol mass from chemical transport models, using satellite data as a constraint. The inverse method effectively makes aerosol forcing a tunable parameter (kludge) for the model, particularly in the presatellite era. Further, key processes associated with the interactions between aerosols and clouds are either neglected or treated with simple parameterizations in climate model simulations evaluated in the AR4.

Given the large uncertainties in forcings and model inadequacies in dealing with these forcings, how is it that each model does a credible job of tracking the twentieth-century global surface temperature anomalies (Fig. 9.5 in Hegerl et al. 2007)? Schwartz (2004) notes that the intermodel spread in modeled temperature trend expressed as a fractional standard deviation is much less than the corresponding spread in either model sensitivity or aerosol forcing, and this comparison does not consider differences in solar and volcanic forcing. This agreement is accomplished through inverse calculations, whereby modeling groups can select the forcing dataset and model parameters that produce the best agreement with observations. While some modeling groups may have conducted bona fide forward calculations without any a posteriori selection of forcing datasets and model parameters to fit the twentieth-century time series of global surface temperature anomalies, the available documentation on each model's tuning procedure and rationale for selecting particular forcing datasets is not generally available.

The inverse calculations can mask variations in sensitivity among the different models. If a model's sensitivity is high, then greater aerosol forcing is used to counter the greenhouse warming, and vice versa for low model sensitivity (Kiehl 2007). Schwartz (2004) argues that uncertainties in aerosol forcing must be reduced at least three-fold for uncertainty in climate sensitivity to be meaningfully reduced and bounded. Further, kludging and neglect of ontic uncertainty in the tuning can result in a model that is over- or undersensitive to certain types or scales of forcing.

With regard to the ability of climate models to simulate natural internal variability on decadal time scales, "there has been little work evaluating the amplitude of Pacific decadal variability in [coupled climate models]" (Randall et al. 2007, p. 621). Whereas most climate models simulate something that resembles the meridional overturning circulation (MOC), the mechanisms "that control the variations in the MOC are fairly different across the ensemble of [coupled climate models]" (p. 621). Comparison of the power spectra of observed and modeled global mean temperatures in Fig. 9.4 of Hegerl et al. (2007) shows that all models underestimate the amplitude of variability on periods of 40–70 yr, which encompasses key modes of multidecadal natural internal variability, such as the Pacific decadal oscillation and the Atlantic multidecadal oscillation.

Bootstrapped plausibility.

If it was so, it might be, and if it were so, it would be; but as it isn't it ain't. That's logic!

—Charles Lutwidge Dodgson (Lewis Carroll)

Bootstrapped plausibility (Agassi 1974) occurs with a proposition that is rendered plausible that in turn lends plausibility to some of the proposition's more doubtful supporting arguments. As such, bootstrapped plausibility occurs in the context of circular reasoning, which is fallacious because of a flawed logical structure whereby the proposition to be proved is implicitly or explicitly assumed in one of the premises. This subsection argues that the IPCC's detection and attribution arguments involve circular reasoning, and that confidence in the evidence and argument is elevated by bootstrapped plausibility.

Consider the following argument that apparently underlies the general reasoning behind the AR4's attribution statement:

1) *Detection.* Climate change in the latter half of the twentieth century is detected based primarily upon increases in global surface temperature

anomalies that are far larger than can be explained by natural internal variability.

2) *Confidence in detection.* The quality of agreement between model simulations with twentieth-century forcing and observations supports the likelihood that models are adequately simulating the magnitude of natural internal variability on decadal to century time scales. From Hegerl et al. 2007, p. 693): "However, models would need to underestimate variability by factors of over two in their standard deviation to nullify detection of greenhouse gases in near-surface temperature data (Tett et al. 2002), which appears unlikely given the quality of agreement between models and observations at global and continental scales (Figs. 9.7 and 9.8) and agreement with inferences on temperature variability from NH temperature reconstructions of the last millennium."
3) *Attribution.* Attribution analyses, including climate model simulations for the twentieth-century climate, that combine natural and anthropogenic forcing agree much better with observations than simulations that include only natural forcing. From Hegerl et al. (2007, p. 684): "The fact that climate models are only able to reproduce observed global mean temperature changes over the 20th century when they include anthropogenic forcings, and that they fail to do so when they exclude anthropogenic forcings, is evidence for the influence of humans on global climate."
4) *Confidence in attribution.* Detection and attribution results based on several models or several forcing histories suggest that the attribution of a human influence on temperature change during the latter half of the twentieth century is a robust result. From Hegerl et al. (2007, p. 669): "Detection and attribution results based on several models or several forcing histories do provide information on the effects of model and forcing uncertainty. Such studies suggest that while model uncertainty is important, key results, such as attribution of a human influence on temperature change during the latter half of the 20th century, are robust."

The strong agreement between forced climate model simulations and observations for the twentieth century (premise 3) provides bootstrapped plausibility to the models and the external forcing data. However, this strong agreement depends heavily on inverse modeling, whereby forcing datasets and/or model parameters are selected based upon the agreement between models and the time series of twentieth-century observations. Further confidence in the models is provided by premise 4, even though the agreement of different models and forcing datasets arises from the selection of forcing datasets and model parameters by inverse calculations designed to agree with the twentieth-century time series of global surface temperature anomalies. This agreement is used to argue that "Detection and attribution studies using such simulations suggest that results are not very sensitive to moderate forcing uncertainties" (Hegerl et al. 2007, p. 678).

Confidence in the climate models that is elevated by inverse calculations and bootstrapped plausibility is used as a central premise in the argument that climate change in the latter half of the twentieth century is much greater than can be explained by natural internal variability (premise 1). Premise 1 underlies the IPCC's assumption (Hegerl et al. 2007, p. 684) that "Global mean and hemispheric-scale temperatures on multi-decadal time scales are largely controlled by external forcings (Stott et al. 2000)" and not natural internal variability. In effect, the IPCC's argument has eliminated multidecadal natural internal variability as a causative factor for twentieth-century climate change. Whereas each model demonstrates some sort of multidecadal variability (which might be of a reasonable amplitude or associated with the appropriate mechanisms), the ensemble averaging process filters out the simulated natural internal variability since there is no temporal synchronization in the simulated chaotic internal oscillations among the different ensemble members.

The IPCC's detection and attribution method is meaningful to the extent that the models agree with observations against which they were not tuned and to the extent that the models agree with each other in terms of attribution mechanisms. The AR4 has demonstrated that greenhouse forcing is a plausible explanation for warming in the latter half of the twentieth century, but it cannot rule out substantial warming from other causes, such as solar forcing and internal multidecadal ocean oscillations owing to the circular reasoning and to the lack of convincing attribution mechanisms for the warming during 1910–40 and the cooling during the 1940s and 1950s.

Bootstrapped plausibility and circular reasoning in detection and attribution arguments can be avoided by the following:

- Using the same best estimate of forcing components from observations or forward modeling for multimodel ensembles
- Conducting tests of the sensitivity to uncertainties associated with the forcing datasets using a single model

- Improving understanding of multidecadal natural internal variability and the models' ability to simulate its magnitude
- Improving detection and attribution schemes to account for the models' inability to simulate the timing of phases of natural internal oscillations and the meridional overturning circulation
- Considering the broad range of confounding factors in assessing likelihood and confidence, including observational errors, model errors and uncertainties, uncertainties in internal variability, and inadequacies in the fingerprinting methodology

The experimental design being undertaken for the Coupled Model Intercomparison Project phase 5 simulations (Taylor et al. 2011) to be used in the IPCC AR5 shows improvements that should eliminate some of the circular reasoning that was evident in the AR4 attribution argument. In the CMIP5 simulations, the use of specific best-estimate datasets of forcing for solar and aerosols is recommended. The National Center for Atmospheric Research (NCAR) Community Climate System Model twentieth-century simulations for CMIP5 (Gent et al. 2011) arguably qualifies as a completely forward calculation, with forcing datasets being selected a priori and no tuning of parameters in the coupled model to the twentieth-century climate other than the sea ice albedo and the low cloud relative humidity threshold. The results of NCAR's CMIP5 calculations show that after 1970, the simulated surface temperature increases faster than the data, so that by 2005 the model anomaly is 0.4°C larger than the observed anomaly. Understanding this disagreement should provide an improved understanding of the model uncertainties and uncertainties in the attribution of the recent warming. This disagreement implies that the detection and attribution argument put forth in the AR4 that was fundamentally based on the good agreement between models and observations will not work in the context of at least some of the CMIP5 simulations.

Since no traceable account is given in the AR4 of how the likelihood assessment in the attribution statement was reached, it is not possible to determine what the qualitative judgments of the lead authors were on the methodological reliability of their claim. Further, the attribution statement itself is at best imprecise and at worst ambiguous: what does "most" mean—51% or 99%? The high likelihood of the imprecise "most" seems rather meaningless (uncertainty monster simplification). From the IAC: "In the Committee's view, assigning probabilities to imprecise statements is not an appropriate way to characterize uncertainty."

Logic of the attribution statement.

Often, the less there is to justify a traditional custom, the harder it is to get rid of it.

—Mark Twain

Over the course of the four IPCC assessments, the attribution statement has evolved in the following way:

- FAR (IPCC 1990, p. xii): "The size of the warming over the last century is broadly consistent with the prediction by climate models, but is also of the same magnitude as natural climate variability . . . Thus the observed increase could be largely due to this natural variability: alternatively this variability and other human factors could have offset a still larger human-induced greenhouse warming. The unequivocal detection of the enhanced greenhouse effect from observations is not likely for a decade or more."
- SAR (IPCC 1995, p. 4): "The balance of evidence suggests a discernible human influence on global climate."
- TAR (IPCC 2001, p. 5): "There is new and stronger evidence that most of the warming observed over the last 50 years is attributable to human activities."
- AR4 (IPCC 2007, p.10): "Most of the observed increase in global average temperatures since the mid-20th century is *very likely* due to the observed increase in anthropogenic greenhouse gas concentrations."

The attribution statements have evolved from "discernible" in the SAR to "most" in the TAR and AR4, demonstrating an apparent progressive exorcism of the uncertainty monster. The attribution statements are qualitative and imprecise in the sense of using words such as "discernible" and "most." The AR4 attribution statement is qualified with a "very likely" likelihood. As stated previously by the IAC, assigning probabilities to imprecise statements is not an appropriate way to characterize uncertainty.

The utility of the IPCC's attribution statement is aptly summarized by this quote from a document discussing climate change and national security (Rogers and Gulledge 2010, p. 19): "For the past 20 years, scientists have been content to ask simply whether ***most*** of the observed warming was caused by human activities. But is the percentage closer to 51 percent or to 99 percent? This question has not generated a great deal of discussion within the scientific community, perhaps because it is not critical to further progress in understanding the climate system. In the policy arena, however, this question is asked often and largely goes unanswered."

The logic of the IPCC AR4 attribution statement is discussed by Curry (2011b). Curry argues that the attribution argument cannot be well formulated in the context of Boolean logic or Bayesian probability. Attribution (natural vs anthropogenic) is a shades-of-gray issue and not a black or white, 0 or 1 issue, or even an issue of probability. Toward taming the attribution uncertainty monster, Curry argues that fuzzy logic provides a better framework for considering attribution, whereby the relative degrees of truth for each attribution mechanism can range in degree between 0 and 1, thereby bypassing the problem of the excluded middle. There is general agreement that the percentages of warming each attributed to natural and anthropogenic causes is less than 100% and greater than 0%. The challenge is to assign likelihood values to the distribution of the different combinations of percentage contributions of natural and anthropogenic contributions. Such a distribution may very well show significant likelihood in the vicinity of 50/50, making a binary demarcation at the imprecise "most" a poor choice.

TAMING THE UNCERTAINTY MONSTER.

I used to be scared of uncertainty; now I get a high out of it.

—Jensen Ackles

Symptoms of an enraged uncertainty monster include increased levels of confusion, ambiguity, discomfort, and doubt. Evidence that the monster is currently enraged includes doubt that was expressed particularly by European policy makers at the climate negotiations in Copenhagen (Van der Sluijs et al. 2010), defeat of a 7-yr effort in the U.S. Senate to pass a climate bill centered on cap and trade, increase in prominence of skeptics in the news media, and the formation of an InterAcademy Independent Review of the IPCC.

The monster is too big to hide, exorcise, or simplify. Increasing concern that scientific dissent is underexposed by the IPCC's consensus approach argues for ascendancy of the monster detection and adaptation approaches. The challenge is to open the scientific debate to a broader range of issues and a plurality of viewpoints and for politicians to justify policy choices in a context of an inherently uncertain knowledge base (e.g., Sarewitz 2004). Some ideas for monster taming strategies at the levels of institutions, individual scientists, and communities are presented.

Taming strategies at the institutional level.

The misuse that is made [in politics] of science distorts, politicizes and perverts that same science, and now we not only must indignantly cry when science falters, we also must search our consciences.

—Diederik Samsom

The politics of expertise describes how expert opinions on science and technology are assimilated into the political process (Fischer 1989). A strategy used by climate policy proponents to counter the strategies of the merchants of doubt (Oreskes and Conway 2010; Schneider and Flannery 2009) has been the establishment of a broad international scientific consensus with high confidence levels, strong appeals to the authority of the consensus relative to opposing viewpoints, and exposure of the motives of skeptics. While this strategy might have been arguably useful, needed, or effective at some earlier point in the debate to counter the politically motivated merchants of doubt, these strategies have enraged the uncertainty monster, particularly since the Climategate e-mails and errors that were found in the IPCC AR4 Working Group II (WGII) report (e.g., Van der Sluijs et al. 2010).

Oppenheimer et al. (2007, p.) remark that "the establishment of consensus by the IPCC is no longer as important to governments as a full exploration of uncertainty." The institutions of climate science, such as the IPCC, the professional societies and scientific journals, national funding agencies, and national and international policy-making bodies, have a key role to play in taming the uncertainty monster. Objectives of taming the monster at the institutional level are to improve the environment for dissent in scientific arguments; to make climate science less political, clarify the political values and visions in play; to expand political debate; and to encourage experts in the social sciences, humanities, and engineering to participate in the evaluation of climate science and its institutions. Identifying areas where there are important uncertainties should provide a target for research funding.

Taming strategies for the individual scientist.

Science . . . never solves a problem without creating ten more.

—George Bernard Shaw

Individual scientists can tame the uncertainty monster by clarifying the confusion and ambiguity associated with knowledge versus ignorance and objectivity versus subjectivity. Morgan et al. (2009) argue that doing a good job of characterizing and dealing with uncertainty can never be reduced to a simple cookbook, and that one must always think critically and continually ask questions. Spiegelhalter (2011) provided the following advice at the recent workshop on Handling Uncertainty in Science at the Royal Society:

- We should try and quantify uncertainty where possible
- All useful uncertainty statements require judgment and are contingent
- We need clear language to honestly communicate deeper uncertainties with due humility and without fear
- For public confidence, trust is more important than certainty

Richard Feynman's (1974, p. 11) address on "cargo cult science" clearly articulates the scientist's responsibility: "Details that could throw doubt on your interpretation must be given, if you know them. You must do the best you can—if you know anything at all wrong, or possibly wrong—to explain it. If you make a theory, for example, and advertise it, or put it out, then you must also put down all the facts that disagree with it, as well as those that agree with it . . . In summary, the idea is to try to give *all* of the information to help others to judge the value of your contribution; not just the information that leads to judgment in one particular direction or another."

Impact of integrity on the monster.

He who fights with monsters might take care lest he thereby become a monster.

—Friedrich Nietzsche

Integrity is an issue of particular importance at the science–policy interface, particularly when the scientific case is represented by a consensus that is largely based on expert opinion. Integrity is to the uncertainty monster as garlic is to a vampire.

Gleick (2011) distinguishes a number of tactics that are threats to the integrity of science: appealing to emotions, making personal (ad hominem) attacks, deliberately mischaracterizing an inconvenient argument, inappropriate generalization, misuse of facts and uncertainties, false appeal to authority, hidden value judgments, selectively omitting inconvenient measurement results, and packing advisory boards.

The issue of integrity is substantially more complicated at the science–policy interface, particularly since the subject of climate change has been so highly politicized. A scientist's statement regarding scientific uncertainty can inadvertently become a political statement that is misused by the merchants of doubt for political gain. Navigating this situation is a considerable challenge, as described by Pielke (2007). Individual scientists can inadvertently compromise their scientific integrity for what they perceive to be good motives. Whereas such actions can provide temporary political advantages or temporarily bolster the influence of an individual scientist, the only remedy in the long run is to let the scientific process take its course and deal with uncertainty in an open and honest way.

The hopeful monster.

There are very few monsters who warrant the fear we have of them.

—Andre Gide

The "hopeful monster" is a colloquial term used in evolutionary biology to describe the production of new major evolutionary groups. Here we invoke the hopeful monster metaphor to address the possibility of taming the monster through the evolution of new entities, enabled by social computing.

When the stakes are high and uncertainties are large, Funtowicz and Ravetz (1993) point out that there is a public demand to participate and assess quality, which they refer to as the extended peer community. The extended peer community consists not only of those with traditional institutional accreditation that are creating the technical work but also those with much broader expertise that are capable of doing quality assessment and control on that work.

New information technology and the open knowledge movement are enabling the hopeful monster. These new technologies facilitate the rapid diffusion of information and sharing of expertise, giving hitherto unrealized power to the peer communities. This newfound power has challenged the politics of expertise, and the "radical implications of the blogosphere" (Ravetz 2010) are just beginning to be understood. Climategate illustrated the importance of the blogosphere as an empowerment of the extended peer community, whereby "criticism and a sense of probity were injected into the system by the extended peer community from the (mainly) external blogosphere" (Ravetz 2010).

While the uncertainty monster will undoubtedly evolve and even grow, it can be tamed through understanding and acknowledgement, and we can learn to live with it by adapting our policies to explicitly include uncertainty. Beck et al.'s (2009, p. 59) statement describes a tamed and happy monster: "Being open about uncertainty should be celebrated: in illuminating where our explanations and predictions can be trusted and in proceeding, then, in the cycle of things, to amending their flaws and blemishes."

ACKNOWLEDGMENTS. This paper was motivated by the Royal Society's Workshop on Handling Uncertainty in Science, held in London, 22–23 March 2010. We would like to thank the denizens of the blog Climate Etc. (www.judithcurry.com) who have provided stimulating discussions and also links to useful reference material on several topics discussed in this paper. We would also like to thank the following individuals who provided comments on the paper: Brian Hoskins, Dan Hughes, Mark Jelinek, Tim Palmer, Roger Pielke Sr., William Rossow, Hans von Storch, Paul Baer, and Michael Welland. The anonymous reviewers provided very helpful comments that improved the paper substantially. This research was supported by NOAA through a subcontract from STG (Curry) and Grant NSF-ATM (Webster).

Editor's Note: Articles published under the Forum heading are intended to spark debate among our readers and in the community. Such dialog is showcased in the Comment (DOI:10.1175/BAMS-D-11-00191.1) and Reply (DOI:10.1175/BAMS-D-11-00195.1) feature that immediately follows this Forum article.

REFERENCES

Agassi, J., 1974: Criteria for plausible arguments. *Mind,* **83,** 406–416.

Annan, J. D., 2010: Bayesian approaches to detection and attribution. *Wiley Interdiscip. Rev.: Climate Change,* **1,** 486–489.

Beck, M. B., and Coauthors, 2009: Grand challenges of the future for environmental modeling: In the setting of NSF's environmental observatories initiatives. White Paper, National Science Foundation, 135 pp. [Available online at www.snre.umich.edu/scavia/wp-content/uploads/2009/12/nsfwhitepaper_beck-07.pdf.]

Bush, V., 1945: Science, the endless frontier: A report to the president. U.S. Government Printing Office, 184 pp. [Available online at www.nsf.gov/od/lpa/nsf50/vbush1945.htm.]

CCSP, 2003: Strategic plan for the U.S. Climate Change Science Program. Final Rep., U.S. Global Change Research Information Office, 202 pp. [Available online at www.climatescience.gov/Library/stratplan2003/final/ccspstratplan2003-all.pdf.]

Curry, J. A., 2011a: Reasoning about climate uncertainty. *Climatic Change,* **108,** 723–732, doi:10.1007/s10584-011-0180-z.

——, 2011b: Nullifying the climate null hypothesis. *Wiley Interdiscip. Rev.: Climate Change,* **2,** 919–924.

Feynman, R. P., 1974: Cargo cult science. California Institute of Technology Engineering and Science Publ., 10–13. [Available online at http://calteches.library.caltech.edu/3043/1/CargoCult.pdf.]

——, 1988: *What Do You Care What Other People Think?* W. W. Norton, 255 pp.

Fischer, F., 1989: *Technocracy and the Politics of Expertise.* Sage Publications, 392 pp.

Forster, P., and Coauthors, 2007: Changes in atmospheric constituents and in radiative forcing. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 129–234.

Funtowicz, S. O., and J. R. Ravetz, 1993: Science for the post-normal age. *Futures,* **25,** 739–755.

Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate,* **24,** 4973–4991.

Gleick, P., cited 2011: Threats to the integrity of science. Testimony to the Senate Committee on Commerce, Science, and Transportation for the hearing on climate change research and scientific integrity. [Available online at www.pacinst.org/publications/testimony/Gleick_Senate_Commerce_2-7-07.pdf].

Gray, L. J., J. D. Haigh, and R. G. Harrison, 2005: The influence of solar changes on the Earth's climate. Hadley Centre Tech. Note 62, 81 pp.

Guillemot, H., 2010: Connections between simulations and observation in climate computer modeling. Scientists' practices and "bottom-up epistemology" lessons. *Stud. Hist. Philos. Mod. Phys.,* **41,** 242–252.

Hargreaves, J. C., 2010: Skill and uncertainty in climate models. *Wiley Interdiscip. Rev.: Climate Change,* **1,** 556–564.

Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 663–745.

IPCC, 1990: Climate Change: The IPCC Scientific Assessment. *Report prepared for Intergovernmental Panel on Climate Change by Working Group I.* Cambridge University Press.

——, 1995: Climate Change 1995: The Science of Climate Change. *Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel and Climate Change.* Cambridge University Press.

——, 2001: Summary for Policymakers. Climate Change 2001: The Physical Science Basis. *Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press.

——, 2007: Summary for Policymakers. Climate Change 2007: The Physical Science Basis. *Contribution of*

Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, S. Solomon et al., Eds., Cambridge University Press.

Johnston, J. S., 2010: *Global Warming Advocacy Science: A Cross Examination.* University of Pennsylvania, Institute for Law and Economics Research Paper 10-08, 82 pp. [Available online at http://ssrn.com/abstract=1612851.]

Judd, K., L. A. Smith, and A. Weisheimer, 2007: How good is an ensemble at capturing truth? Using bounding boxes for forecast evaluation. *Quart. J. Roy. Meteor. Soc.,* **133,** 1309–1325.

Kandlikar, M., J. Risbey, and S. Dessai, 2005: Representing and communicating deep uncertainty in climate change assessment. *C. R. Geosci.,* **337,** 443–455.

Kiehl, J. T., 2007: Twentieth century climate model response and sensitivity. *Geophys. Res. Lett.,* **34,** L22710, doi:10.1029/2007GL031383.

Knutti, R., 2008: Should we believe model predictions of future climate change? *Philos. Trans. Roy. Soc. London,* **A366,** 4647–4664.

——, and Coauthors, 2008: A review of uncertainties in global temperature projections over the twenty-first century. *J. Climate,* **21,** 2651–2663.

Lenhard, J., and E. Winsberg, 2011: Holism and entrenchment in climate model validation. *Science in the Context of Application: Methodological Change, Conceptual Transformation, Cultural Reorientation,* M. Carrier and A. Nordmann, Eds., Boston Studies in the Philosophy of Science Series, Vol. 274, Springer, 115–130.

McWilliams, J., 2007: Irreducible imprecision in atmospheric and oceanic simulations. *Proc. Natl. Acad. Sci. USA,* **104,** 8709–8713.

Mitchell, J. F. B., and Coauthors, 2001: Detection of climate change and attribution of causes. *Climate Change: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 695–738.

Morgan, G., H. Dowlatabadi, M. Henrion, D. Keith, R. Lempert, S. McBride, M. Small, and T. Wilbanks, Eds., 2009: Best practice approaches for characterizing, communicating, and incorporating scientific uncertainty in climate decisionmaking. Final Rep., U.S. Climate Change Science Program Synthesis and Assessment Product 5.2., 94 pp. [Available online at www.climatescience.gov/Library/sap/sap5-2/final-report/default.htm.]

Moss, R. H., and S. H. Schneider, 2000: Uncertainties in the IPCC TAR: Recommendations to lead authors for more consistent assessment and reporting. Guidance papers on the cross cutting issues of the Third Assessment Report of the IPCC, R. Pachauri, T. Taniguchi, and K. Tanaka, Eds., IPCC, 33–51.

Muller, P., 2010: Constructing climate knowledge with computer models. *Wiley Interdiscip. Rev.: Climate Change,* **1,** 565–580.

Oppenheimer, M., B. C. O'Neill, M. Webster, and S. Agrawala, 2007: The limits of consensus. *Science,* **317,** 1505–1506.

Oreskes, N., 2007: The scientific consensus on climate change: How do we know we're not wrong? *Climate Change: What It Means for Us, Our Children, and Our Grandchildren,* J. F. C. DiMento and P. Doughman, Eds., MIT Press, 65–99.

——, and E. M. Conway, 2010: *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming.* Bloomsbury Press, 368 pp.

Palmer, T. N., 2001: A nonlinear dynamical perspective on model error: A proposal for nonlocal stochastic-dynamic parameterization in weather and climate prediction models. *Quart. J. Roy. Meteor. Soc.,* **127,** 279–304.

——, G. J. Shutts, R. Hagedorn, F. J. Doblas-Reyes, T. Jung, and M. Leutbecher, 2005: Representing model uncertainty in weather and climate prediction. *Annu. Rev. Earth Planet. Sci.,* **33,** 163–193.

Parker, W. S., 2009: Confirmation and adequacy-for-purpose in climate modeling. *Aristotelian Soc. Suppl. Vol.,* **83,** 233–249.

——, 2010: Predicting weather and climate: Uncertainty, ensembles and probability. *Stud. Hist. Philos. Mod. Phys.,* **41,** 263–272.

Petersen, A. C., 2006: *Simulating Nature: A Philosophical Study of Computer-Simulation Uncertainties and Their Role in Climate Science and Policy Advice.* Het Spinhuis, 210 pp.

Pielke, R., Jr., 2007: *The Honest Broker: Making Sense of Science in Policy and Practice.* Cambridge University Press, 188 pp.

Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 589–662.

Ravetz, J., cited 2010: Climategate: Plausibility and the blogosphere in the post-normal age. [Available online at http://wattsupwiththat.com/2010/02/09/climategate-plausibility-and-the-blogosphere-in-the-post-normal-age/.]

Rogers, W., and J. Gulledge, 2010: Lost in translation: Closing the gap between climate science and national security policy. Center for a New American Security Rep., 58. [Available online at www.cnas.org/files/documents/publications/Lost%20in%20Translation_Code406_Web_0.pdf.]

Sarewitz, D., 2004: How science makes environmental controversies worse. *Environ. Sci. Policy,* **7,** 385–403.

Schneider, S. H., and T. Flannery, 2009: *Science as a Contact Sport: Inside the Battle to Save Earth's Climate.* National Geographic Society, 304 pp.

Schwartz, S. E., 2004: Uncertainty requirements in radiative forcing of climate change. *J. Air Waste Manage. Assoc.,* **54,** 1351–1359.

Shukla, J., R. Hagedorn, B. Hoskins, J. Kinter, J. Marotzke, M. Miller, T. N. Palmer, and J. Slingo, 2009: Revolution in climate prediction is both necessary and possible: A declaration at the World Modelling Summit for Climate Prediction. *Bull. Amer. Meteor. Soc.,* **90,** 175–178.

Smith, L. A., 2002: What might we learn from climate forecasts? *Proc. Natl. Acad. Sci. USA,* **99,** 2487–2492.

——, 2006: Predictability past predictability present. *Predictability of Weather and Climate,* T. N. Palmer and R. Hagedorn, Eds., Cambridge University Press, 217–250.

Spiegelhalter, D., cited 2011: Quantifying uncertainty. [Available online http://downloads.royalsociety.org/audio/DM/DM2010_03/Speigelhalter.mp3.]

Stainforth, D. A., M. R. Allen, E. R. Tredger, and L. A. Smith, 2007: Confidence, uncertainty and decision-support relevance in climate predictions. *Philos. Trans. Roy. Soc. London,* **A365,** 2145–2161.

Stott, P. A., S. F. B. Tett, G. S. Jones, M. R. Allen, J. F. B. Mitchell, and G. J. Jenkins, 2000: External control of 20th century temperature by natural and anthropogenic forcings. *Science,* **290,** 2133–2137.

Swart, R., L. Bernstein, M. Ha-Duong, and A. C. Petersen, 2009: Agreeing to disagree: Uncertainty management in assessing climate change, impacts and responses by the IPCC. *Climatic Change,* **92,** 1–29.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2011: A summary of the CMIP5 experiment design. CLIVAR, 33 pp. [Available online at www.clivar.org/organization/wgcm/references/Taylor_CMIP5.pdf.]

Tett, S. F. B., and Coauthors, 2002: Estimation of natural and anthropogenic contributions to twentieth century temperature change. *J. Geophys. Res.,* **107,** 4306, doi:10.1029/2000JD000028.

U.S. DOD, cited 2011: DOD news briefing—Secretary Rumsfeld and Gen. Myers. [Available online at www.defense.gov/transcripts/transcript.aspx?transcriptid=2636.]

Van der Sluijs, J. P., 2005: Uncertainty as a monster in the science–policy interface: Four coping strategies. *Water Sci. Technol.,* **52,** 87–92.

——, R. van Est, and M. Riphagen, Eds., 2010: Room for climate debate: Perspectives on the interaction between climate politics, science and the media. Rathenau Instituut, 96 pp. [Available online at www.rathenau.nl/uploads/tx_tferathenau/Room_for_climate_debate.pdf.]

Walker, W. E., P. Harremoes, J. Rotmans, J. P. van der Sluijs, M. B. A. van Asselt, P. Janssen, and M. P. Krayer von Krauss, 2003: Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integr. Assess.,* **4,** 5–17.

Webster, M., 2009: Uncertainty and the IPCC: An editorial comment. *Climatic Change,* **92,** 37–40.

WMO, 2005: Guidance notes for lead authors of the IPCC Fourth Assessment Report on addressing uncertainties. IPCC, 4 pp. [Available online at www.ipcc.ch/pdf/supporting-material/uncertainty-guidance-note.pdf.]