Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences

Naomi Oreskes,* Kristin Shrader-Frechette, Kenneth Belitz

Verification and validation of numerical models of natural systems is impossible. This is because natural systems are never closed and because model results are always nonunique. Models can be confirmed by the demonstration of agreement between observation and prediction, but confirmation is inherently partial. Complete confirmation is logically precluded by the fallacy of affirming the consequent and by incomplete access to natural phenomena. Models can only be evaluated in relative terms, and their predictive value is always open to question. The primary value of models is heuristic.

In recent years, there has been a dramatic increase in the use of numerical simulation models in the earth sciences as a means to evaluate large-scale or complex physical processes. In some cases, the predictions generated by these models are considered as a basis for public policy decisions: Global circulation models are being used to predict the behavior of the Earth's climate in response to increased CO_2 concentrations; resource estimation models are being used to predict petroleum reserves in ecologically sensitive areas; and hydrological and geochemical models are being used to predict the behavior of toxic and radioactive contaminants in proposed waste disposal sites. Government regulators and agencies may be required by law to establish the trustworthiness of models used to determine policy or to attest to public safety (*1*, *2*); scientists may wish to test the veracity of models used in their investigations. As a result, the notion has emerged that numerical models can be "verified" or "validated," and techniques have been developed for this purpose (*1*, *3–5*). Claims about verification and validation of model results are now routinely found in published literature (*6*).

Are claims of validity and verity of numerical models legitimate (*2*, *7*)? In this article, we examine the philosophical basis of the terms "verification" and "validation" as applied to numerical simulation models in the earth sciences, using examples from hydrology and geochemistry. Because demand for the assessment of accuracy in numerical modeling is most evident at the interface between public policy and scientific usage, we focus on examples relevant to policy (*8*). The principles illustrated, however, are generic.

N. Oreskes is in the Department of Earth Sciences and the Department of History, Dartmouth College, Hanover, NH 03755. K. Shrader-Frechette is in the Department of Philosophy, University of South Florida, Tampa, FL 33620. K. Belitz is in the Department of Earth Sciences, Dartmouth College, Hanover, NH 03755.

*To whom correspondence should be addressed.

Verification: The Problem of "Truth"

The word verify (from Latin, *verus*, meaning true) means an assertion or establishment of truth (*9*). To say that a model is verified is to say that its truth has been demonstrated, which implies its reliability as a basis for decision-making. However, it is impossible to demonstrate the truth of any proposition, except in a closed system. This conclusion derives directly from the laws of symbolic logic. Given a proposition of the form "p" entails "q," we know that if "p" is true, then "q" is true if and only if the system that this formalism represents is closed.

For example, I say, "If it rains tomorrow, I will stay home and revise this paper." The next day it rains, but you find that I am not home. Your verification has failed. You conclude that my original statement was false. But in fact, it was my intention to stay home and work on my paper. The formulation was a true statement of my intent. Later, you find that I left the house because my mother died, and you realize that my original formulation was not false, but incomplete. It did not allow for the possibility of extenuating circumstances (*10*). Your attempt at verification failed because the system was not closed.

This example is trivial, but even an apparently trivial proposition can be part of a complex open system. Indeed, it is difficult to come up with verbal examples of closed systems because only purely formal logical structures, such as proofs in symbolic logic and mathematics, can be shown to represent closed systems. Purely formal structures are verifiable because they can be proved by symbolic manipulations, and the meaning of these symbols is fixed and not contingent on empirically based input parameters (*11*).

Numerical models may contain closed mathematical components that may be verifiable, just as an algorithm within a computer program may be verifiable (*12*). Mathematical components are subject to verification because they are part of closed systems that include claims that are always true as a function of the meanings assigned to the specific symbols used to express them (*13*). However, the models that use these components are never closed systems. One reason they are never closed is that models require input parameters that are incompletely known. For example, hydrogeological models require distributed parameters such as hydraulic conductivity, porosity, storage coefficient, and dispersivity, which are always characterized by incomplete data sets. Geochemical models require thermodynamic and kinetic data that are incompletely or only approximately known. Incompleteness is also introduced when continuum theory is used to represent natural systems. Continuum mechanics necessarily entails a loss of information at the scale lower than the averaging scale. For example, the Darcian velocity of a porous medium is never identical to the velocity structure at the pore scale. Finer scale structure and process are lost from consideration, a loss that is inherent in the continuum mechanics approach.

Another problem arises from the scaling-up of nonadditive properties. The construction of a numerical simulation model of a ground-water flow system involves the specification of input parameters at some chosen scale. Typically, the scale of the model elements is on the order of meters, tens of meters, or kilometers. In contrast, the scale on which input parameters are measured is typically much smaller, and the relation between those measurements and larger scale model parameters is always uncertain and generally unknown. In some cases, it is possible to obtain input data at the scale chosen by the modeler for the model elements (for example, pump tests), but this is not often done, for practical reasons. Even when such measurements are available, they are never available for all model elements (*14*).

Another reason hydrological and geochemical models are never closed systems is that the observation and measurement of both independent and dependent variables are laden with inferences and assumptions. For example, a common assumption in many geochemical models of water-rock interaction is that observable mineral as-

useful advantages for interdisciplinary researchers. First, systems models provide a way to codify knowledge from different disciplines into a unified and coherent framework. Second, they encourage focused and disciplined thinking about the causal relationships in a system. Third, they allow researchers, ecosystem managers, and stakeholders to explore how their system may respond to a variety of scenarios so that responses can be formulated and management actions can be implemented. However, system models can only achieve these advantages if they are developed and used deliberately and thoughtfully.

Developing simulation models is part science and part craft; there are no general, infallible rules. Different practitioners process their experiences in different ways. In this paper, we offer 10 heuristics for interdisciplinary modeling that we have developed over a period of several years through our experiences in a variety of integrated research projects. The primary audience we have in mind is people who are not presently engaged in interdisciplinary research but are interested in moving in this direction in the future. However, we also hope to stimulate thinking, discussing, and writing about methodology among current modeling practitioners, and we believe that our emphasis on rapid prototyping and sensitivity analysis will be of interest to them.

What do we mean by "heuristic"? Polya (1945) defined this term as "the name of a certain branch of study" whose aim is "to understand the methods and rules of discovery and invention." However, in this essay, the word is used in the sense defined by Starfield and others (1994): "a heuristic is a plausible or reasonable approach that has often proved to be useful, a rule of thumb."

In other words, this is a paper in which the findings have been generated inductively from our collective experiences on a range of interdisciplinary projects rather than a deductive literature review that investigates the success or failure of other projects based on whether they did or did not use these heuristics.

To illustrate our 10 heuristics, we give examples of lessons we have learned from developing integrated interdisciplinary models for a recent project investigating the Sustainability of Arctic Communities (SAC). This project involved a team of 25 scientists (representing eight different disciplines in both the natural and the social sciences) and residents from four indigenous Arctic communities in the Yukon Territory, Northwest Territories, and Alaska. Research team members came from several universities and from government agencies. The goal of the project was to investigate how changes in climate, tourism, oil development, and government funding could affect the sustainability of our partner communities. The communities themselves defined their goals for sustainability in the early part of the project (G.P. Kofinas and others unpublished). These goals included (a) maintaining a strong relationship with the land and the animals, (b) developing healthy mixed economies (that is, a subsistence harvesting economy in parallel with a cash economy), (c) exercising local control over land use and resource use in their homelands, (d) educating their young people in both traditional knowledge and Western science while also educating outsiders about their way of life; and (e) maintaining a thriving native culture (evidenced, for example, by the use of indigenous language, respect for community elders, and spending time on the land). In other words, the communities saw sustainability not simply in terms of sustainable resource use, but also in economic, political, and sociocultural terms. To address this holistic set of community goals, it was obviously essential to take an interdisciplinary view of the system. An integrated approach was in any case implicit in the framing of the original project proposal and in the range of disciplinary scientists included in the research team. Their expertise covered the fields of vegetation ecology, caribou biology, caribou behavior, household economies, cultural ecology, social anthropology, policy analysis, Arctic tourism, and natural resource modeling.

The emphasis of this paper is not on the SAC Project itself, although examples will be drawn from that project to illustrate our heuristics. Also, the heuristics given here relate primarily to scientists working with other scientists on interdisciplinary projects rather than to scientists working with stakeholding. The SAC Project not only served to bring scientists together, but also involved residents of indigenous Arctic communities. A companion paper to this one (G.P. Kofinas and others unpublished) offers heuristics for researcher–stakeholder interactions and for synthesizing local knowledge and science. Finally, although we discuss various aspects of teamwork and collaboration, our focus is not on collaboration generally (as in, for example, Gray 1985, 1991 or Kofinas and Griggs 1996) but on the process of the collaborative development of synthesis models.

Heuristic 1. Know what skills to look for when recruiting an interdisciplinary team. It is not a foregone conclusion that any given team of specialists will work together effectively to produce a tightly integrated view of a system. Indeed, there are many challenges and obstacles that must be addressed

before a variety of scientists can work together effectively in an interdisciplinary mode. Among these obstacles is the problem of cross-discipline communication, since specialists are used to interacting with peers from within their fields who share a common view of the issues and a common language for discussing them. The problem of communication will be addressed in heuristic 9, but it is also relevant here because it can be a stumbling block when choosing and recruiting team members. To foster good communication among prospective team members even at the recruitment stage, it is essential to develop a prototype conceptual model of the system (see heuristic 3).

A major obstacle to interdisciplinary work is that scientists are trained and socialized from their graduate school days to focus on narrow, tractable problems within clearly defined boundaries. They are taught how to identify problems that lie on the cutting edge of their discipline, and they learn appropriate methods for solving these problems. In other words, using Holling's (1996) distinction between the science of parts and the science of the integration of parts, scientific training is essentially an induction into the methods and norms of the science of *parts*. In the science of parts, investigators within a discipline focus on a narrowly defined question with the goal of reducing uncertainty to the point of consensus. In contrast, the science of the *integration of parts* calls for people who are committed to studying a complex system by focusing not so much on the individual components of the system as on the interrelationships among its components. Although it is often true that outstanding interdisciplinarians also have very high reputations within a specialist field, the best disciplinary minds are not necessarily the best interdisciplinary team members. Interdisciplinary projects are intellectually demanding in a different way from classic reductionist science, and they need at least some big-picture researchers who will creatively explore the linkages and interfaces between their own discipline and other fields of inquiry in which they may themselves have no special expertise.

Another key attribute of a good interdisciplinary team member is the ability to simplify what is known and, when necessary, guess at the unknown (see heuristic 8). These activities call for people with a deep grasp of their own disciplines: Weak or insecure disciplinary minds can frustrate team progress by refusing to explore linkages, to simplify their field, or to guess at unknown factors. Not only are these activities necessary to make an interdisciplinary study a success, but they are often not the types of activities that their own disciplinary peers will recognize as valuable contributions to the scholarship of their field, and publishing their work may not be easy. Young scientists are particularly at risk because they have not yet established their reputation and because the reward systems of academia bend to favor disciplinary specialists.

All team members who embark on interdisciplinary projects need to be made aware of these kinds of problems in advance so that they join the team with realistic expectations, an adventurous attitude, and a willingness to work at cross-disciplinary communication. For a project leader to know if someone is right for the team, he or she should look for scholars who can see the big picture, whose track record shows an ability to work with people outside their own discipline, who are good listeners, and whose interest in a problem outweighs their concern for career advancement! How can such people be motivated to participate? One incentive may simply be an appeal the intellectual satisfaction of seeing how their disciplinary interests fit within a larger framework, thereby sharpening their understanding of their own disciplines.

Heuristic 2. Invest strongly in problem definition early in the project. By their very nature, projects that involve complex systems with many interacting components lend themselves to multiple focuses. Different stakeholders often have different perceptions of the problem. In addition, each disciplinary expert has a vested professional interest in defining the problem so as to give his or her discipline a prominent role with a bias toward the researcher's particular expertise within the discipline. In the absence of strong leadership, it is far easier to end up doing multidisciplinary research (where experts work in parallel with each other without much meaningful integration) than it is to do truly interdisciplinary research. To address this challenge, the problem needs to be thoughtfully and clearly defined from the outset. This is never a straightforward exercise, even when there are compelling reasons for the study. The parties involved in problem definition need to understand that choosing the focus of an IA project is fundamentally a negotiated process. For this reason, all the parties (disciplinary researchers, stakeholders, and funding agencies) must be given opportunities to exchange perspectives and must be aware of each other's priorities.

Heuristics 1 and 2 are obviously interrelated. Until you define the problem, you cannot assemble a team; and until you have a team, you cannot really define the problem. (This is why we promoted the idea of developing a prototype conceptual model, even at the stage of recruiting team members.) The ideal situation is one in which a small group has the

opportunity to make an initial attempt to define the problem and then go on to recruit the additional expertise necessary. This kind of opportunity requires either project development funding or an infrastructure that brings the small group together.

In the Sustainability of Arctic Communities study, the High Latitude Ecosystems Directorate (HLED) of the US Man and the Biosphere (MAB) program provided an opportunity for a group of natural and social scientists to interact with each other across disciplines and to formulate a rough and preliminary project definition. The HLED group of six individuals started discussions 2 years before the funding opportunity arose and decided to focus on the combined effects of future climate change and oil development on barren ground caribou (*Rangifer tarandus*) and the indigenous communities that depend on caribou as a subsistence resource. At this initial stage, the stakeholder communities were not directly involved. This was a mistake, even though several group scientists had worked with indigenous communities for many years and therefore had a good grasp of the issues involved. On the basis of the preliminary problem definition, the group obtained permission from the US MAB committee to advertise position descriptions for additional HLED members with expertise in cultural ecology, modeling, and caribou biology. As new people were recruited, they brought new perspectives on the problem, and the process of negotiating a common focus continued. The stakeholder communities joined the project in the 1st year and provided an important reality check on our understanding of the issues (G.P. Kofinas and others unpublished).

It is extremely difficult to anticipate the appropriate problem definition at the outset of a project. In fact, during the early part of the SAC project, the participating scientists felt that the target was forever shifting. In hindsight, the team ought to have built more structure into the negotiation process to ensure convergence on the problem definition. To promote clarity of thought and allow the group members to see whether the initial problem definition is correct, we recommend the following structured procedures:

1. Cooperate on the development of first prototype "straw" system simulation models, and
2. Submit the current understanding of the system to a "peer review" by stakeholders. The discipline of having to articulate the problem definition to an audience beyond the team itself helps to get the ideas clear.

Heuristic 3. Use rapid prototyping for all modeling efforts. Not only is it hard to define the problem correctly on the first attempt, it is also extremely difficult at the start of a new project to discern the relative importance of each of the components. Therefore, rapid prototyping of models is essential. Instead of trying to specify at the outset of the project precisely what the final model will look like and what questions it will address, the participants should recognize that the first year will be devoted to the development of a prototype model aimed at clarifying the objectives of the study. Moreover, in subsequent years, the problem and the model will be further refined through successive prototypes (see Schrage 2000 for a number of case studies from the business world in which prototyping and successive refinement consistently led to superior final products). It is only when project participants see actual output from the model that they can begin to grasp the big picture and gain an understanding of the system dynamics as a whole. This understanding allows them to place their own contributions in perspective. Furthermore, it is only when a prototype model is up and running that the relative importance of the various components of the system or weaknesses in the framing of the original hypotheses gradually begin to emerge.

For example, the original proposal for the SAC Project envisioned a model time horizon of 100–200 years. The revised proposal (with added emphasis on the Arctic communities) defined a time horizon of 40 years. One of the original hypotheses was that climate change would lead to changes in summer vegetation biomass and plant community composition, that caribou herds would be affected by these changes, and that Arctic communities, in turn, would be impacted by the caribou. On the basis of this definition of the problem, a set of sustainability indicators was developed; these included plant biomass, caribou herd size, hunters' time-on-the-land, and seasonal caribou harvest. We designed a synthesis model that would address the problem as it had been defined. However, in the process of developing and testing the model, we discovered that there are time lags of 50–100 years before any substantial simulated effects of climate change are apparent at a plant community or biomass level (Epstein and others 2000); within a 40-year time horizon, climate-related vegetation changes were therefore almost insignificant. We also learned from the initial modeling exercise that as long as the caribou herd size is above a certain threshold (estimated to be about 60% of its present level), annual caribou migration patterns affect harvest success far more than a decline or increase in

herd size. This suggests that the initial emphasis on herd population dynamics may have been somewhat misplaced. In both of these examples, our problem definition led us to believe that certain factors were more important than they turned out to be.

It is also important to make sure that the linkages among different parts of the system are strong and that the system behavior is not dominated by a single component (in which case the problem does not necessarily call for an interdisciplinary approach). Our experience, and that of other practitioners (for example, Holling 1978; Walters 1986), has shown that it is more fruitful to begin with the system itself and to look outward to the components rather than to look piecemeal at the system from within the perspective of the individual components.

One danger in interdisciplinary modeling work is that people who are not fluent in systems modeling may not engage properly with the task. The solution is not to recruit only model-oriented scientists (which would limit the scope and breadth of the synthesis), but rather to work at drawing nonmodelers into the process. Prototype models that are simple enough to demonstrate and explain to all team members are an essential step in the education of nonmodelers.

Heuristic 4. Allow the project's focus to evolve by not allocating all funds up front. This is a luxury seldom available to research scientists, given the current policy of multiyear, multi-investigator projects. However, one of the inherent difficulties with interdisciplinary research is that defining the problem often represents a major part of the project. Thus, a chicken-and-egg situation arises. The problem cannot be defined until a working team is in place, but it is impossible to know how deeply to involve specific team members until the problem has been defined. Even when the problem is apparently well defined, it is extremely hard to assess a priori which components determine the system dynamics most strongly until a first prototype of the synthesis work has been constructed. It is likely that the relative importance of the various components will only emerge during the study. We have already alluded to the initial hypothesis of climate change → vegetation change → caribou herd dynamics → caribou availability to human communities. By the time we discovered that this apparently central hypothesis was not a main driver of change, the project's funds had been allocated and could not easily be shifted to address newly evolving hypotheses.

It might be better if funding agencies awarded preliminary planning funds (say, for the 1st year or through the development of a first prototype model) and then funded the remainder of the project only when it was demonstrated that the correct mix of scientists was working together effectively and attacking a well-defined problem. If all the funds are committed up front for the full duration of the study, the project leadership has no flexibility to add new people as their expertise becomes necessary or to reallocate funds from a component of the work that offers little to the integrated effort.

Heuristic 5. Ban all models or model components that are inscrutable. An "inscrutable" model is a black box in which the inner workings are inaccessible to all but the original developers. The user is required to take the output on faith. The problem with inscrutable models is that people have no incentive to engage with them intellectually. If the model produces any counterintuitive results, people cannot access the logic that led to those results. It is not surprising then that their usual reaction is to lose trust in the model rather than ask about the intermediate relationships that led to those final results.

In the SAC Project, a complex model of caribou energetics (Hovey and others 1989; Kremsater 1991; Daniel 1993) was initially thought to be essential at the interface between vegetation change and caribou population dynamics. We realized later that what we really needed were models of herd distribution and movement, but a commitment had already been made to this energetics model. Until we developed a much simpler caribou population model, the project depended on output from a black box model that only a few people understood and used. A top-down, rapid-prototyping approach could have helped avoid this situation. Graphical "box and arrow" representations of the system (Jørgensen 1986; Walters 1986) combined with the simplest possible component models, programmed using software that is easily accessible to all team members (such as spreadsheets), allow a team of scientists from different disciplinary backgrounds to understand and engage with the key relationships of the model.

Heuristic 6. Instead of concentrating on one all-purpose synthesis model, invest in a suite of models, each with a well-defined objective. This heuristic applies particularly to the collaborative development stage of a project. It allows participants from a subset of disciplines to engage with models that focus on the interfaces between those subsets.

The idea of building a suite of models may seem to go against the very idea of interdisciplinary synthesis modeling, but meshing existing submodels together can be a difficult and time-consuming ex-

ercise. Submodels may operate at different time scales because of the nature of the underlying processes. Similar variables in two submodels may be represented at different levels of detail from one another. The probabilistic outcomes produced by one stochastic submodel may not translate easily into hard-and-fast input values for other deterministic submodels in the system.

Although it is essential to represent adequately the logic and the results of each submodel in all the other relevant submodels to which it links, if this is done properly, it may not be necessary to have one "supersynthesis" model that runs each submodel within the same overall programming framework. In fact, for quality-control purposes, it is probably good to have some kind of human interface between submodels. This allows the results of each submodel to be assessed and the quality of its conclusions evaluated, so it can be determined how best to include the insights gained from each model in the next submodel. This process helps to determine which details are not essential and allows the development of a "boiled down" whole system model at a level of abstractism that may initially have been unacceptable to some team members.

A further argument for a suite of models is that few users of the overall synthesis model will be interested in all of its components. Most people have an interest in only three or four of the outcomes of the model. A suite of models allows users to examine the components with which they are familiar and to see how these results fit with the outcomes they expect, based on their knowledge and experience of that part of the system. However, for nonscientific users, it is helpful to have a seamless interface that allows them to explore whichever part of the system they are most interested in. If no such interface exists, users will not readily recognize that they are seeing an integrated view of the system, and much of the benefit of the exercise will be lost.

Heuristic 7. Maintain a healthy balance between the well-understood and the poorly understood components of the system. All system models are balancing acts between what one knows and understands and what one does not know. The temptation is to put too much emphasis on those parts of the system where understanding and data are good and to ignore or gloss over the areas where little is known. This is not surprising, given the way in which the scientific enterprise tends to favor specialists. For example, even though there may be a clear and obvious link between caribou migration and household economic production, a caribou biologist might know a great deal about caribou activities and energetics, but relatively little about herd migration patterns. An economist may have a good understanding of the factors that influence people's decisions to take wage employment, but know relatively little about the factors that account for the successful harvest and production of caribou meat on the land itself.

Furthermore, people like to concentrate on the details they know about and understand (Likens 1998). In particular, scientists are socialized into an epistemological framework that places a high value on detailed quantitative hard facts and tends to take a dim view of uncertainty (even when the uncertainty involves an educated guess in an important area where little else is known). They are often skeptical of simplification and even more uncomfortable with the idea of making educated guesses. For example, the village economy model we developed for the SAC Project contained over 90 different job categories, each characterized in terms of its required education level, its seasonal availability, and whether men or women were more likely to be found in that position. These definitions were firmly grounded in survey data. However, the economists on our team were reluctant to speculate on how these definitions might evolve over the next 40 years; as a consequence, the model contained the implicit assumption that social norms such as gender preferences for job types would not change during two generations.

Maintaining the balance between the known and the unknown requires strong project leadership. In a review of IA projects, Parson (1995) observed: "Since researchers working within their fields do not normally attend to borders of other fields, achieving this attention shift requires some form of authority in an assessment project, or at least a coordinating mechanism and a common language for communicating across boundaries." One way to achieve such coordination would be to bring in an outside modeling consultant who could facilitate key workshops. In addition to providing a fresh viewpoint, an outsider who has the trust of the team could also provide the kind of authority that Parson refers to. Modelers are certainly not the only people who could fulfill this role. The two most important qualifications are an ability to see the big picture and the earned trust and respect of other team members. These qualities may well be present in one of the disciplinary specialists if he or she is also a good big-picture scholar. However, through rapid prototyping and sensitivity analysis, modeling can be particularly useful for ranking the relative importance of the parts and processes in the model, as well as making a rapid assessment of the value

and differences between alternative conjectures about the unknowns. The synthesis modelers should be encouraged and empowered to use their skills to help resolve the tensions between simplicity and detail that are inherent to any modeling project (Costanza and Sklar 1985; Starfield and Bleloch 1991).

Heuristic 8. Sensitivity analysis is vital at all stages of the modeling effort. Thorough sensitivity analysis involves testing not only different parameter values but also the assumptions and the effect of alternative educated guesses at the underlying processes (see, for example, Starfield and others 1995; Starfield and Bleloch 1991). Sensitivity analysis is the only available means of determining what goes into the model and what level of detail is necessary. It is an essential tool for estimating the likely effects of alternative hypotheses for system processes. Sensitivity analysis should not simply be thought of as an automated process that tests all parameters, but rather an important part of the culture of modeling that is used for the thoughtful exploration of alternative assumptions. It follows that the work of sensitivity analysis should be done by most (ideally, all) of the project team, not just the modeler. Because each person on the team brings a different perspective to the problem, he or she is thus likely to run different experiments and uncover different problems. In fact, team efforts are essential both for identifying implicit assumptions (social norms do not change during two generations, for example) and for developing plausible alternative scenarios as part of the sensitivity analysis.

Sensitivity tests are essentially mini-experiments. To be effective in shaping the prototype modeling process, the models supporting the mini-experiments need to run virtually in real time. Waiting days, weeks, or months for model results is too long. We found that the ability to work as a group to set up a model simulation, and then view the results within a minute or two, was principally responsible for most advances in developing model relationships that crossed disciplinary boundaries.

Heuristic 9. Work hard at communication and budget for face-to-face meetings. Effective communication lies at the heart of interdisciplinary research. Not only is it necessary for scientists to engage with one another to produce an integrated view of a system, their findings must also be explained clearly to stakeholders and to the public. In the Internet era, communication can take many forms, including list-server memos, e-mails, phone calls, small face-to-face work groups, plenary team meetings, and public meetings. Each communication medium serves a different purpose, and it is dangerous to assume that simply because we have these tools at our disposal, people from different disciplinary backgrounds will automatically communicate effectively with each other. In the SAC Project, scientists often appeared to have reached a point of understanding in their discussions, only to find out later that in fact they had two rather different things in mind (sometimes as the result of using the same words but meaning different things by them). Team members need to make an effort to become more familiar with each other's mental frameworks and to be cognizant of what specific people mean when they use certain words or concepts. This is one reason why rapid prototyping is so valuable. It leads quickly to a product that provides a common language and enables participants to say "No, that's not really what I have in mind."

One way to foster better communication is by developing simulation models in easily accessible modeling environments, such as spreadsheets. The goal is to work continually toward a culture of transparent and accessible models, so as to ensure that the models are understandable to everyone on the team. In this regard, we have found that spreadsheets have several advantages over traditional programming languages such as FORTRAN, BASIC, or C++. Most scientists are familiar with the spreadsheet environment and its basic concepts. Also, spreadsheets allow us to quickly and easily develop straw models as part of the dialogue among the participants, so we can constantly point to something tangible and ask, "Is this what you mean?" The built-in graphing functions of spreadsheets enable graphical model output with very little programming effort. Finally, because spreadsheets perform calculations each time a cell is changed, they are powerful tools for sensitivity analysis.

In addition to communication within the team, a second area requiring careful consideration is how the integrated work of the team will be communicated to the stakeholders and the public, a task that is vastly underrepresented in many scientific projects. Interdisciplinary research that affects people's lives directly must be explained to them in accessible language, stripped of its technical scientific terminology. The results need to be put into everyday terms, and it is crucial to spell out both the practical implications of the findings and the areas of uncertainty. Funding agencies need to be willing to support the outreach and extension part of interdisciplinary research, and scientists need the help of communication specialists to get their results into public discourse in a form that can be digested and discussed. The SAC Project's efforts at outreach include the development of a simplified interactive

Web-based model interface that we call the "Possible Futures Model" (G.P. Kofinas and others unpublished). The model demonstrates our attempts at ongoing innovation in all areas of the interface between model and user: ease of use, hypertext documentation, graphical output, and built-in features for explaining model results and documenting users' feedback comments.

An important lesson from the SAC Project is that the budget for face-to-face meetings was inadequate (both for meetings among researchers and meetings between researchers and community partners). The proposal did include funds for annual project meetings of the entire team; these were of some value, but we found it far more profitable to hold work sessions involving small numbers of researchers from the various components to develop specific component linkages. E-mail is not an effective medium for planning and for the creative generation of ideas. Written exchanges work best once there is a common understanding of the problem, common assumptions, and a negotiated set of task assignments; face-to-face meetings are indispensable for these groundwork decisions. Because face-to-face meetings are so much richer in communicative content and because they allow trust to be built more easily than can be done in a series of written messages (Daft and Huber 1987), meetings with component researchers are also critical to the work of the synthesis modeler. When we began to hold these meetings, considerable momentum was gained. Face-to-face contact should be a nonnegotiable part of any IA budget, particularly when team members are geographically dispersed.

Heuristic 10. Approach the project with humility. Even though the scientists on the team may be world-class experts in their respective component fields, they are all likely to be amateurs when it comes to the system as a whole. It is worth remembering that a distinguished group of component experts does not guarantee a distinguished *system* team. In fact, since laypeople often have a deep and holistic understanding of their local environment, we scientists may be no more "expert" than they are, even though their knowledge is not necessarily scientific. All team members must take the time to probe and query each other's approaches, assumptions, and methods. More important, they must be willing to have their own assumptions and statements probed by others. This requires humility, a willingness to be challenged by team members outside one's own area, and an openness to learning from such transactions. The excitement and challenge of interdisciplinary research lies in uncovering together the unknown—namely, the behavior of the system.

Humility and caution are especially important when scientists work on projects that are intended to inform policy, thereby affecting people's lives. Synthesis modelers bear the brunt of the responsibility of ensuring, first, that the assumptions behind the models are carefully spelled out and, second, that robust conclusions are shown to be robust even in the face of uncertainty.

In the spirit of humility, we acknowledge that the 10 heuristics presented here are obviously not exhaustive. Integrated assessment is a very complicated business, and as a form of inquiry, it is still in the early stages of development. Not only are the dynamics of large complex systems hard to understand, but the challenge of bringing disparate perspectives together is a formidable one. We offer these principles simply because we believe it is important for synthesis modelers and interdisciplinarians alike to reflect on what they have done, in the hope of doing it better the next time around.

ACKNOWLEDGMENTS

Financial support was provided by the National Science Foundation (NSF OPP-95-21459). We acknowledge the role played by members of the Sustainability of Arctic Communities research team and the communities of Old Crow, Aklavik, Fort McPherson, and Arctic Village in our collaborative effort to build a shared understanding of a complex system. Ann Kinzig and an anonymous reviewer offered helpful comments on a previous version of the manuscript.

REFERENCES

Costanza R, Sklar FH. 1985. Articulation, accuracy and effectiveness of mathematical models: a review of freshwater wetland applications. Ecol Model 27:45–68.

Daft RL, Huber GP. 1987. How organizations learn: a communication framework. Sociol Org 5:1–36.

Daniel CJ, 1993. Computer simulation models for the Porcupine Caribou Herd: user's guide. Richmond Hill: (Ontario): ESSA Ltd., for Canadian Wildlife Service, Environment Canada, Whitehorse, Yukon, 17 p.

Epstein HE, Walker M, Chapin FS, Starfield AM, 2000. A transient, nutrient-based model of arctic plant community response to climatic warming. Ecol Appl 10:824–41.

Gray B. 1991. Collaborating: finding common ground for multiparty problems. San Francisco: Jossey-Bass.

Gray B. 1985. Conditions facilitating inter-organizational collaboration. Hum Rel 38(10):911–36.

Gunderson LH, Holling CS, Light SS, editors. 1995. Barriers and bridges to the renewal of ecosystems and institutions. New York: Columbia University Press. 593 p.

Holling CS. 1978. Adaptive environmental assessment and management. London: Wiley.

Holling CS. 1996. Surprise for science, resilience for ecosystems and incentives for people. Ecol Appl 6(3):733–5.

Holling CS. 1998. Two cultures of ecology. Conserv Ecol [online] 2(2):4. Available on the Internet: URL: http://www.consecol-.org/vol2/iss2/art4.

Holling CS. 1995. What barriers? What bridges? In: Gunderson LH, Holling CS, Light SS, editors. Barriers and bridges to the renewal of ecosystems and institutions. New York: Columbia University Press. p 3–34.

Hovey FW, Kremsater LL, White RG, Russell DE, Bunnell FL, 1989. Computer simulation models of the Porcupine Caribou Herd. II. Growth. Technical report no. 54. Canadian Wildlife Service, Pacific and Yukon Region, Vancouver B.C.

Jørgensen SE. 1986. Fundamentals of ecological modeling. Amsterdam: Elsevier. 389 p.

Kofinas GP, Braund SR, Archie B, Charlie J Sr, Eamer J, James S, Kruse K, Nicolson C, Tetlichi J. Integrating local knowledge in an integrated assessment: lessons from the Sustainability of Arctic Communities Project experience. Unpublished.

Kofinas GP, Griggs JR. 1996. Collaboration and the B.C. Round Table on the Environment and the Economy: an analysis of a "better way" of deciding. Environments 23(2):17–40.

Kremsater LL. 1991. Brief description of computer simulation models of the Porcupine Caribou Herd. In: Butler CE, Mahoney ST, editors. Proceedings of the 4th North American Caribou Workshop. St. Johns (Newfoundland): p 299–314.

Likens GE. 1998. Limitations to intellectual progress in ecosystem science. In: Pace M, Groffman P, editors. Successes, limitations and frontiers in ecosystem science. New York: Springer-Verlag. p 247–71.

Meadows DH, Meadows DL, Randers J, Behrems WW. 1972. Limits to growth, New York: Universe Books. 205 p.

Nissani M. 1997. Ten cheers for interdisciplinarity: the case for interdisciplinary knowledge and research. Soc Sci J 34(2): 210–6.

Parson E. 1995. Integrated assessment and environmental policy making. Energy Policy 23(4/5):463–75.

Polya G. 1945. How to solve it. Princeton (NJ): Princeton University Press. 253 p.

Riley GA. 1947. A theoretical analysis of the zooplankton population of Georges Bank. J Mar Res 6(2):104–13. (Quoted in: Swartzman GL, Kaluzny SP. 1987. Ecological simulation primer. New York: Macmillan. 370 p).

Risbey J, Kandlikar M, Dowlatabadi H. 1996. Learning from integrated assessment of climate change. Clim Change 34: 369–95.

Schrage M. 2000. Serious play: how the world's best companies simulate to innovate. Boston: Harvard Business School Press. 244 p.

Starfield AM, Bleloch AL. 1991. Building models for conservation and wildlife management. 2nd ed. Edina (MN): Interaction. 252 p.

Starfield AM, Roth JD, Ralls K. 1995. "Mobbing" in Hawaiian monk seals (*Monachus schauinslani*): the value of simulation modeling in the absence of apparently crucial data. Conserv Biol 9(1):166–74.

Starfield AM, Smith KA, Bleloch AL. 1994. How to model it: problem solving for the computer age. 2nd ed. Edina (MN): Interaction. 206 p.

Walters CJ. 1986. Adaptive management of renewable resources. New York: Macmillan. 374 p.