SCIENCE

Published by the **American Association for the Advancement of Science (AAAS)**, *Science* serves its readers as a forum for the presentation and discussion of important issues related to the advancement of science, including the presentation of minority or conflicting points of view, rather than by publishing only material on which a consensus has been reached. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by the AAAS or the institutions with which the authors are affiliated.

The American Association for the Advancement of Science was founded in 1848 and incorporated in 1874. Its objectives are to further the work of scientists, to facilitate cooperation among them, to foster scientific freedom and responsibility, to improve the effectiveness of science in the promotion of human welfare, to advance education in science, and to increase public understanding and appreciation of the importance and promise of the methods of science in human progress.

Membership/Circulation
Director: Michael Spinella
Deputy Director: Marlene Zendell
Member Services: Rebecca Dickerson, *Manager*; Mary Curry, *Supervisor*; Pat Butler, Helen Williams, Laurie Baker, *Representatives*
Marketing: Dee Valencia, *Manager*; Jane Pennington, *Europe Manager*; Hilary Baar, *Associate*; Angela Mumeka, *Coordinator*
Business and Finance: Jacquelyn Roberts, *Manager*; Robert Smariga, *Assistant Manager*
Administrative Assistant: Nina Araujo de Kobes
***Science* Member Services**
Marion, Ohio: 800-347-6969;
Washington, DC: 202-326-6417
Other AAAS Programs: 202-326-6400

Advertising and Finance
Associate Publisher: Beth Rosner
Advertising Sales Manager: Susan A. Meredith
Recruitment Advertising Manager: Janis Crowley
Advertising Business Manager: Deborah Rivera-Wienhold
Finance: Randy Yi, *Senior Analyst;* Shawn Williams, *Analyst*
Marketing: John Meyers, *Manager;* Allison Pritchard, *Associate*
Traffic Manager: Tina Turano
Recruitment: Terri Seiter, *Assistant Manager*; Dan Moran, *Traffic Manager*; Debbie Cummings, Celeste Wakefield, Angela Wheeler, *Sales*
Reprints Manager: Corrine Harris
Permissions Manager: Arlene Ennis
Sales Associate: Carol Maddox

ADVERTISING SALES: East Coast/E. Canada: Richard Teeling, 201-904-9774, FAX 201-904-9701 • **Southeast:** Mark Anderson, 305-856-8567, FAX 305-856-1056 • **Midwest:** Elizabeth Mosko, 312-665-1150, FAX 312-665-2129 • **West Coast/W. Canada:** Neil Boylan, 415-673-9265, FAX 415-673-9267 • **UK, Scandinavia, France, Italy, Belgium, Netherlands:** Andrew Davies, (44) 457-838-519, FAX (44) 457-838-898 • **Germany/Switzerland/Austria:** Tracey Peers, (44) 270-760-108, FAX (44) 270-759-597 • **Japan:** Mashy Yoshikawa, (3) 3235-5961, FAX (3) 3235-5852
Recruitment: 202-326-6555, FAX 202-682-0816
European Recruitment: AnneMarie Vis, (44) 0223-302067, FAX (44) 0223-302068
Australia/New Zealand Recruitment: Keith Sandell, (61) 02-922-2977, FAX (61) 02-922-1100
Send materials to *Science* Advertising, 1333 H Street, NW, Washington, DC 20005.

Information for Contributors appears on pages 37–39 of the 7 January 1994 issue. Editorial correspondence, including requests for permission to reprint and reprint orders, should be sent to 1333 H Street, NW, Washington, DC 20005.
Internet addresses: science_editors@aaas.org (for general editorial queries); science_letters@aaas.org (for letters to the editor); science_reviews@aaas.org (for returning manuscript reviews)

LETTERS

Ice Age "Venuses"

The Grimaldi Ice Age figurines referred to in Random Samples of 18 February (p. 923) were shown in 1914 to George McCurdy of the Peabody Museum at Harvard University by the daughter of the excavator, L. Jullien. In 1939 they were offered for sale to the American Museum of Natural History, who sent photographs to the Peabody Museum. Hallam Movius bought the most important of these for the Peabody. It is important, not because it is a "masterpiece," but because it is the only Ice Age carving depicting *two* females, one side pregnant, and the other not, thus confirming the presence of both concepts in Ice Age female imagery (Fig. 1). The long archaeological argument as to whether pregnancy and nonpregnancy were depicted appears to have been settled by my analysis of the figurine published in 1986 (*1*). The remaining Grimaldi figurines are patently "crude" carvings. It was likely for this reason that they were not purchased earlier.

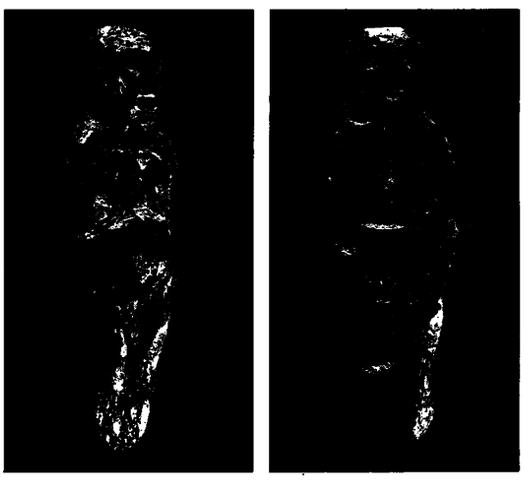


Fig. 1. Grimaldi steatite figurine, three-quarter (left) and rear (right) view, from (*1*, p. 809) (Courtesy A. Marshack).

Early in the 20th century, museums and collectors were primarily interested in purchasing obvious "masterpieces." The best of the Grimaldi figurines had been purchased early in the century for the Musée des Antiquitées Nationales in France. Such "masterpieces" had long shaped archaeological theories concerning the quality and meaning of the so-called "Venus" figurines.

The new set of Grimaldi figurines may, therefore, be important precisely because of their ostensible "poor" quality. There are many classes and types of female imagery with different apparent meanings and uses, and one class, patently "crude" and often relatively quickly made (but not works in progress), was apparently produced by less skilled carvers for short-term, or even one-time, ritual use (*2*). The "masterpieces," on the other hand, often show evidence of long-term curation, use, and reuse.

Alexander Marshack
Peabody Museum,
Harvard University,
Cambridge, MA 02138, USA

References

1. A. Marshack *L'Anthropologie* **90**, 807 (1986).
2. ———, *Proc. Prehist. Soc.* **57**, 17 (1991); *The Roots of Civilization* (Moyer-Bell, Mount Kisko, NY, 1991), p. 282.

I have read with some amusement the speculations concerning the purposes of the 20,000-year-old "Venus figurines." It seems to me that the simplest and most obvious explanation for them has been overlooked, namely, that they are the Epipaleolithic equivalent of the centerfold.

L. R. Caswell
Department of Chemistry,
Texas Women's University,
Denton, TX 76204–1973, USA

■

The Meaning of Models

Three cheers for Naomi Oreskes, Kristin Shrader-Frechette, and Kenneth Belitz for their much needed critique of "verification, validation, and confirmation of numerical models . . ." (Articles, 4 Feb., p. 641).

Two points require amplification. First, the impossibility of "verifying" or "validating" models is not limited to computer models. All theories invoke auxiliary hypotheses and rely on imperfectly measured data. Any nontautological theory (that is, any theory that refers to the world) is underdetermined and thus unverifiable, whether it is embodied in a large-scale computer model or consists of the simplest equations. The differences between traditional, analytic theories and large simulations are differences of degree only. Large-scale models, because they require estimation of many more parameters and boundary conditions, have been the focus of the

debate over validation, but the limits on the certainty of our knowledge described in the article apply to all.

Second, as the authors note, the principles discussed apply to models outside the earth sciences, including economics and the social sciences. Although positivism may have "collapsed resoundingly in the 1950s" in philosophy and physics, it continued to wash over the social sciences, cresting with Milton Friedman's still influential 1953 *Essays in Positive Economics* (*1*). Despite numerous critiques (*2*), positivist empiricism remains the leading theory of knowledge in economics today. One need only substitute "demand elasticity" for "porosity" and "marginal propensity to consume" for "hydraulic conductivity" and the article would describe well the state of model assessment in economics. Economic models are routinely described as "being valid" or "having been validated," and the principal (and often only) criterion for "validity" is the correspondence of simulated and actual data. The concept of data, furthermore, is usually restricted to constructs for which numerical measures are available, aggravating the problem of underdetermination. Worse, because it is time consuming to manipulate model inputs until one gets the result one desires, leading economic forecasters have long resorted to "add factoring"—the practice of adding a fudge factor to the *output* of the model so that it corresponds better to the modeler's intuition, thus avoiding the troublesome bother of actually using the model to reach conclusions from well-documented assumptions (*3*). In one study (*4*), the General Accounting Office found that a leading econometric forecasting firm add-factored interest rate projections in a scenario of restricted monetary growth from the model's output of 34% per year to the more sensible value of 7% per year. Such practices persist in part through the failure of model consumers (academics, policymakers, managers, and citizens at large) to look behind the "tests" of "validity" offered by the model makers.

John D. Sterman
Alfred P. Sloan School of Management,
Massachusetts Institute of Technology,
Cambridge, MA 02142, USA

References

1. M. Friedman, Ed., *Essays in Positive Economics* (Univ. of Chicago Press, Chicago, IL, 1953).
2. H. A. Simon, *Am. Econ. Rev.* **69**, 493 (1979); P. A. Samuelson, *ibid.* **53**, 231 (1963); W. Leontieff, *ibid.* **61**, 1 (1971); E. H. Phelps-Brown, *Econ. J.* **82**, 4 (1972); E. Leamer, *Am. Econ. Rev.* **73**, 31 (1983).
3. J. D. Sterman, in *Foresight and National Decisions*, L. Grant, Ed. (University Press of America, Lanham, MD, 1988), pp. 133–169.
4. *Three Large Scale Model Simulations of Four Money Growth Scenarios* (Joint Economic Committee, Subcommittee on Monetary and Fiscal Policy, 97th Cong., 2nd Sess., Washington, DC, 1982).

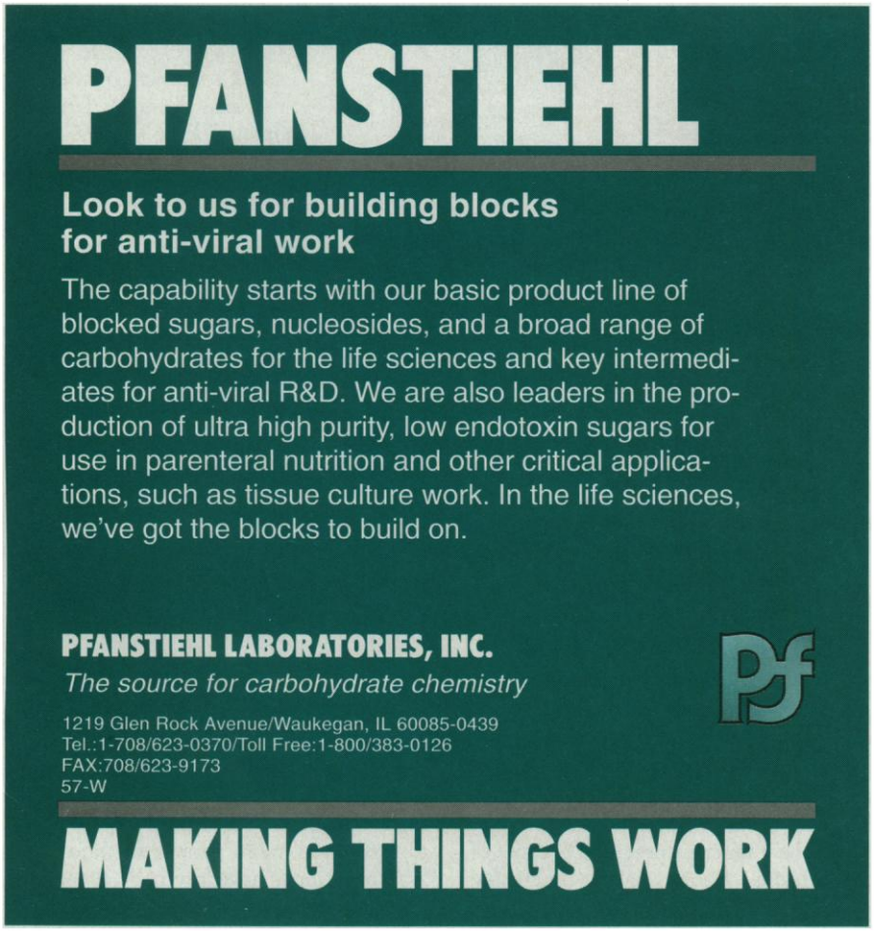
Oreskes *et al.* correctly point out some of the limitations of models and the ways in which models can be tested. However, they set up and convincingly slay a straw man. The straw man is the claim that verification and validation prove that the scientific content of a model is "true." With possible rare exceptions, I know of no case where a modeler has made this claim. "Validation" amounts to acceptance testing. If a model meets specified performance criteria, it is accepted as a "credible representation of the real system" (*1*), and perhaps further as the best technology currently available. That is all that "validation" means in the technical jargon of simulation modeling. For example, a definition of "validation" is "substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model" (*2*).

A complex computer model is not reducible to a simple syllogism. However, the correct syllogism in this case is, If this model meets the specified criteria (p), then it is acceptable for use in the problem domain (q). Validation is testing to demonstrate that the criteria are met, thus asserting p and validly concluding q (the model is acceptable for use). This is not a case of affirming the consequent.

Oreskes *et al.* do not define what they mean by "truth" presumably because we all know what that is. However, they use the term "truth" in in the special context of formal logic, and it has other meanings that are consistent with the technical meanings of "verification" and "validation" (*3*).

Semantically, there is little to choose among the terms "verify," "validate," "confirm," "authenticate," "corroborate," and "substantiate," because they are synonyms in ordinary language (*3*). The term "confirmation" has no more claim to probability or provisional acceptance than any of the others. Thus, "verification" and "validation" (and "confirmation") acquire special disciplinary meanings for testing simulation models.

Modelers themselves should take the lead in asserting the restrictions and limitations of models and should draw some important lessons from Oreskes *et al.*: (i) make clear that "verification" and "validation" are used in a technical sense; (ii) if necessary, don't use the terms if they are likely to be misunderstood and create a false sense of truth rather than consensus; (iii)



Circle No. 9 on Readers' Service Card

take care to specify the context of the model; and (iv) use model acceptability and performance indices rather than simple declarations of validation to describe the results of model testing.

Edward J. Rykiel Jr.
Department of Industrial Engineering, Texas A&M University, College Station, TX 77843–3131, USA

References

1. S. V. Hoover, and R. R. Perry, *Simulation* (Addison-Wesley, Reading, MA 1989).
2. R. G. Sargent, in *Proceedings of the 1984 Winter Simulation Conference*, S. Sheppard, U. Pooch, D. Pegden, Eds. (IEEE 84CH2098-2, Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1984), pp. 115–121.
3. *Webster's New Collegiate Dictionary* (Merriam, Springfield, MA, 1975).

Response: We fully agree with Sterman that the points raised in our article are limited neither to computer models nor to the earth sciences. We focused on numerical models in the earth sciences because the more general point about underdetermination of scientific theories has been made eloquently elsewhere, because earth science is the area of our own experience and expertise, and because the issues of verification and validation are active topics of discussion among earth scientists.

Rykiel sees little semantic difference between "verify," "validate," and "confirm." While acknowledging overlap in the many meanings of these terms, we disagree that they are synonyms in common usage. For example, one can *verify* that a parking permit has been *validated*. Nuances of meaning do matter, particularly when terms are shared in scientific and lay discourse. Our discussion of the terms "verify" and "validate" is an accurate representation of the way many earth scientists use these terms (*1*), and our use of the term "confirm" follows decades of scholarship in logic and in philosophy of science (*2*). Nevertheless, the primary objective of our article was substantive, not semantic. If modelers were to change only language and not practice, then our article would not have achieved its objective. Current usage *is* misleading and can create a false sense of truth, particularly in practical policy applications.

Rykiel says we slay a straw man, but the "straw people" in this case include the International Atomic Energy Agency and the U.S. Department of Energy. The terminology that we critique comes directly from the scientific guidelines of these agencies and from published scientific literature (*3*). The syllogism that Rykiel puts forward as the correct logical construct for evaluating models begs the fundamental questions at stake: Who decides what the specified criteria are? What are the limits of the problem domain? and Can they change with time?

Rykiel concludes that "modelers themselves should take the lead in asserting the restrictions and limitations of models." Insofar as two of us are modelers (*4*), and all three of us routinely use and evaluate models (*5*), this is precisely what we tried to do.

Naomi Oreskes
Kenneth Belitz
Department of Earth Sciences, Dartmouth College, Hanover, NH 03755, USA
Kristin Shrader-Frechette
Environmental Sciences and Policy Program and Department of Philosophy, University of South Florida, Tampa, FL 33620, USA

References and Notes

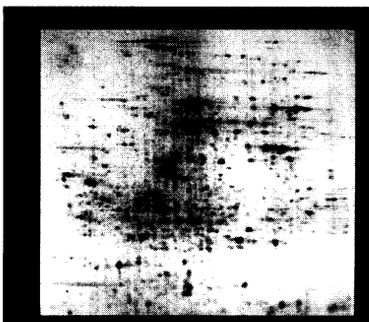
1. For example, J. L. Younker *et al.*, "Report of early site suitability evaluation of the potential repository site at Yucca Mountain, Nevada" (SAIC-91/8000 (U.S. Department of Energy, Washington, DC, 1992).
2. C. G. Hempel and P. Oppenheim, *Philos. Sci.* **15**, 135 (1948); C. G. Hempel, *Aspects of Scientific Explanations* (Free Press, New York, 1965); C. G. Hempel, *Philosophy of Natural Science* (Prentice-Hall, Englewood Cliffs, NJ, 1966); C. Glymour, in *The Philosophy of Science*, R. Boyd, P. Gasper, J. D. Trout, Eds. (Massachusetts Institute of Technology Press, Cambridge, MA, 1991), pp. 485–500.
3. "Radioactive waste management glossary" (IAEA-TECDOC-264, International Atomic Energy Agency, Vienna, 1982); *Environmental Assessment: Yucca Mountain Site, Nevada Research and Development Area, Nevada*, vol. 2, *DOE/RW-0073* (Office of Civilian Radioactive Waste Management, U.S. Department of Energy, Washington, DC, 1986). Also see *U.S. Nuclear Regulatory Commission Radioactive Waste Management Program (NUREG-0865)* (U.S. Nuclear Regulatory Commission, Washington, DC, 1990); P. A. Davis, N. E. Olague, M. T. Goodrich, "Approaches for the validation of models used for performance assessment of high-level nuclear waste repositories" (SAND90-0575/NUREG CR-5537) (Sandia National Laboratories, Albuquerque, NM, 1991); J. L. Younker *et al.* (*1*); *Adv. Wat. Resour.* **15** (1992).
4. K. Belitz, J. M. Gronberg, S. P. Phillips, *U.S. Geological Survey Water Supply Paper 2396* (1993); K. Belitz and J. D. Bredehoeft, in *Hydrogeology of Low-Permeability Environments Hydrogeology, Selected Papers*, vol. 2, S. P. Neuman and I. Neretnieks, Eds. (Verlag-Heinz Heise, Hannover, Germany, 1990) pp. 7–17; K. Belitz and J. D. Bredehoeft, *Am. Assoc. Petrol. Geol. Bull.* **72**, 1334 (1988); N. Oreskes, *Proceedings of the Eighth International Association on the Genesis of Ore Deposits*, p. A137 (July 1990).
5. K. S. Shrader-Frechette, *Burying Uncertainty: Risk and the Case Against Geological Disposal of Nuclear Waste* (Univ. of California Press, Berkeley and Los Angeles, CA, 1993).

Corrections and Clarifications

In the Random Sample "Venuses reappear" (18 Feb., p. 923), Patricia Rice of West Virginia University is incorrectly identified as "Patricia White." Randall White of New York University is the source of quotes attributed to "White."

MILLIPORE
ELECTROPHORESIS
NEW!
Pre-cast pH Stabilized IEF Tube Gels

Improve the performance of isoelectric focusing for 2-D electrophoresis with Millipore pH stabilized IEF gels. The linear 3 - 10 pH gradient is stabilized in the gel and offers better resolution in the acidic and basic regions than can be achieved with standard ampholyte gels.



Wild type yeast sample using Millipore pH stabilized pre-cast IEF gel.

Millipore pH stabilized IEF gels are conveniently pre-cast in patented threaded glass tubes. Ready-to-use, they don't require rehydrating, cooling or messy oil overlays. And, they are compatible with any system that will accept a 1 mm x 26 cm tube gel format.

To introduce you to the performance and convenience of Millipore pre-cast IEF gels, we're offering a two-for-one special. Buy one package of 6 and get a second package free. But act now. This offer expires 9/30/94.

For more information call 1-800-MILLIPORE and ask for product number S1ER 007 E4 or our fax on demand information retrieval system at 1-800-MILLIFX and request #1704. In Japan, call (81) (3) 3471-8191. In Europe, contact the local office or fax us in France (33) (1) 30.12.7189.

MILLIPORE

Circle No. 30 on Readers' Service Card