DISCUSSION:

REVISIONS OF BOOTSTRAP TESTING*

CLARK GLYMOUR

Department of History and Philosophy of Science
University of Pittsburgh

Scientific arguments often attribute the credit or blame for particular experimental or observational results to fragments of theories rather than to the whole of them. The thesis of *Theory and Evidence* (Glymour 1980) is that there are structural criteria that account for at least part of the distribution of credit and blame; the book also attempts a tentative formal characterization of one structural criterion. David Christensen (1983) has produced a series of counterexamples which show that the formal characterization is untenable. The question remains whether the failure is due to an oversight in the formal theory or to the falsity of the very idea that there are structural criteria for evidential relevance. Since the counterexamples are uniformly eliminated by a single formal principle, I incline to the first answer.

Christensen's counterexamples are as follows:

1. E: Ra & Ba
   H_1: $Ax(Rx \rightarrow Bx)$
   H_2: $AxGx$
   E confirms H_2 with respect to H_1 & H_2

$$\begin{array}{ccc} & Gx & \\ / & & \backslash \\ Rx & & Bx \end{array} \qquad Ax((Rx \rightarrow Bx) \leftrightarrow Gx))$$

2. E: Ra & Ba & Sa
   H_1: $Ax(Rx \rightarrow Bx)$
   H_2: $Ax(Sx \rightarrow Cx)$
   E confirms H_2 with respect to H_1 & H_2

$$\begin{array}{ccc} & Cx & \\ / & | & \backslash \\ Rx & Bx & Sx \end{array} \qquad Ax((Rx \rightarrow Bx) \leftrightarrow (Sx \rightarrow Cx))$$

*Received July 1983.

Philosophy of Science, 50 (1983) pp. 626–629.

3. (Kepler's Laws)
   E: $L_1a \;\&\; L_2a \;\&\; Ka \;\&\; -\,(a = b)$
   H_1: AxL_1x
   H_2: AxL_2x
   H_3: $AxAy(Kx = Ky)$
   E confirms H_3 with respect to the conjunction of all three hypotheses.

$$\begin{array}{ccc} & Ky & \\ / & | & \backslash \\ L_1a & L_2b & Ka \;\&\; -\,(a = b) \end{array} \qquad AxAy((L_1x \;\&\; L_2x) \leftrightarrow (Kx = Ky))$$

4. E: $Ra \;\&\; Fa$
   H_1: $Ax(Rx \rightarrow Bx)$
   H_2: $Ax(Rx \rightarrow Fx)$
   E confirms H_1 with respect to H_1 & H_2.

$$\begin{array}{ccc} & Bx & \\ / & & \backslash \\ Rx & & Fx \end{array} \qquad Ax(Rx \rightarrow (Fx \leftrightarrow Bx))$$

In each case the sentence written to the side of the graph is the hypothesis used to compute a value for the quantity at the top of the graph. I have taken modest liberties in representing Christensen's test of Kepler's laws. So represented, each of Christensen's examples violates the following restriction:

R: For all i, H must not entail that the hypothesis Ti used in computing a quantity Qi, occurring essentially in H, is equivalent to an hypothesis Ri whose essential vocabulary is a proper subset of the essential vocabulary of Ti.

Restriction R says in effect that the computations must restrict the quantities occurring essentially in H in a way that is independent of the restriction that H itself imposes on its quantities. In example 1, the restriction is violated because

$$AxGx \vdash Ax((Rx \rightarrow Bx) \leftrightarrow Gx) \leftrightarrow Ax(Rx \rightarrow Bx).$$

In example 2, R is violated because

$$Ax(Sx \rightarrow Cx) \vdash Ax((Rx \rightarrow Bx) \leftrightarrow (Sx \rightarrow Cx)) \leftrightarrow (Rx \rightarrow Bx).$$

Similar violations are readily shown for the remaining examples.

Since the restriction seems intuitively sensible, and since it eliminates all of Christensen's counterexamples, as well as analogous counterexamples for the theory applied to systems of equations, I would like to revise the conditions for bootstrap testing to incorporate it. Doing so requires that the notion of a "computation" be modified.

Let a quantity be as in *Theory and Evidence*; that is, any atomic formula containing no individual constants is a quantity. A value for a quantity is any sentence obtained by substituting constants or definite descriptions for the variables in the quantity. Two quantities are identified if and only if they have the same set of values. Let H be an hypothesis, Qi a quantity occurring essentially in H, and let E be a collection of singular evidence sentences. A computation of Qi from E with respect to theory T is a sentence Ti among the consequences of T that determines a value of Qi from E (that is, E and Ti together entail a value for Qi).

Definition. E bootstrap confirms H with respect to T if and only if

1. E, H and T are jointly consistent;
2. For each quantity Qi occurring essentially in H there exists a computation, Ti, of Qi from E with respect to T;
3. The set of computed values of Qi confirms H according to the satisfaction* condition of *Theory and Evidence*.
4. For all i, H does not entail that Ti is equivalent to any sentence S whose essential vocabulary is properly included in the essential vocabulary of Ti;
5. There exist values for the quantities occurring essentially in E such that the values Qi computed from these values disconfirm H according to the satisfaction* condition.
6. For each i, there is no proper logical consequence Ti^* of Ti such that Ti^* and the remaining Tj satisfy 2 through 5.

Alternative, but related, conceptions of hypothesis testing can be obtained by replacing the satisfaction* condition of clauses 3 and 5, which is essentially Hempel's theory of confirmation generalized to the predicate calculus with identity, by any well defined confirmation relation. There are, furthermore, several alternatives to clause 5, including the following:

5a. The conjunction of the Ti does not entail that H is equivalent to any hypothesis K whose essential vocabulary is properly contained in the essential vocabulary of H.

5b. If K is any sentence in the essential vocabulary of H, and the conjunction of the Ti entails K, then K is valid.

Substituting either 5a or 5b for 5 in the definition above results in alternative formal schemes with slightly different properties. I have no hypothesis as to which of these schemes is the most satisfactory. The three schemes correspond closely to analogous schemes for the testing of equations which I have described in Glymour (1983).

These schemes eliminate the counterexamples Christensen cites, although they may of course be liable to other difficulties. One difficulty to which they are evidently liable is that they prohibit the confirmation of theoretical hypotheses consisting of a universally quantified conjunction. Thus

$$AxMx$$

cannot be bootstrap confirmed by any theory with respect to any evidence in which M is not among the evidential predicates. A similar limitation applies to the analogous treatment of systems of equations. I regard this result as unfortunate but not intolerable. It is better for a formal confirmation theory to be narrow minded than for it to be gullible, and in this case the narrow-mindedness will, I think, only rarely prohibit one from representing plausible arguments about the confirmation of real hypotheses.

REFERENCES

Christensen, D. (1983), "Glymour on Evidential Relevance", *Philosophy of Science 50*: 471–481.

Glymour, C. (1980), *Theory and Evidence*. Princeton, New Jersey: Princeton University Press.

Glymour, C. (1983), "On Testing and Evidence", in John Earman (ed.), *Minnesota Studies in the Philosophy of Science X*. Minneapolis: University of Minnesota Press.