



CHICAGO JOURNALS



The Irrelevance of Bootstrapping

Author(s): David Christensen

Source: *Philosophy of Science*, Vol. 57, No. 4 (Dec., 1990), pp. 644-662

Published by: [The University of Chicago Press](#) on behalf of the [Philosophy of Science Association](#)

Stable URL: <http://www.jstor.org/stable/187766>

Accessed: 03/10/2011 16:06

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press and Philosophy of Science Association are collaborating with JSTOR to digitize, preserve and extend access to *Philosophy of Science*.

<http://www.jstor.org>

THE IRRELEVANCE OF BOOTSTRAPPING*

DAVID CHRISTENSEN†

*Department of Philosophy
University of Vermont*

The main appeal of the currently popular “bootstrap” account of confirmation developed by Clark Glymour is that it seems to provide an account of evidential relevance. This account has, however, had severe problems; and Glymour has revised his original account in an attempt to solve them. I argue that this attempt fails completely, and that any similar modifications must also fail. If the problems can be solved, it will only be by radical revisions which involve jettisoning bootstrapping’s basic approach to theories. Finally, I argue that there is little reason to think that even such drastic modifications will lead to a satisfactory account of relevance.

The “bootstrap” account of confirmation developed in recent years by Clark Glymour (1980, 1983a) has generated a great deal of interest. Bootstrapping attempts to provide an appealing alternative to traditional hypothetico-deductive (H-D) accounts of confirmation: like H-D accounts, it analyzes confirmation in terms of simple deductive relations among the hypotheses of a theory; unlike H-D accounts, bootstrapping offers an account of evidential relevance—a way of determining which *parts* of a theory a given bit of evidence confirms or disconfirms.

Unfortunately, this account of evidential relevance—the centerpiece of the bootstrap account—has proven quite problematic. A series of examples in Christensen (1983) showed that the account, as presented in *Theory and Evidence* (Glymour 1980), did not work. I went on to argue there that the problem seemed to stem not from the details of the account, but from the basic approach bootstrapping shares with classical hypothetico-deductivism: determining evidential relevance relations by looking only at the logical structure of the set of consequences entailed by a theory’s natural first-order axiomatization. I suggested that this basic approach to theories—for convenience, I’ll call this the “classical” approach—would have to be abandoned before an adequate account of evidential relevance could be given.

In response to these problems, Glymour (1983a) has proposed a mod-

*Received June 1988; revised December 1988.

†This paper benefited from discussions or correspondence with James R. Christensen, Clark Glymour, Thomas R. Grimes, Richard Healey, Arthur Kuflik, William E. Mann, and Derk Pereboom.

ified version of the bootstrap account which, however, remains faithful to the classical conception of theories. The modifications are designed to disallow some of the above-mentioned counterexamples; in particular, the modified account will not allow certain clear cases of irrelevant confirmation allowed by the original account. While Glymour admits that his modifications disallow some intuitively correct confirmations,¹ he considers this “unfortunate but tolerable”, since

[i]t 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. (Glymour 1983a, p. 629)

Thus Glymour concludes that it was the details of the original formulation of bootstrapping that were defective, and that the counterexamples do not indicate “the falsity of the very idea that there are structural criteria for evidential relevance”.

Now I am not sure that I understand Glymour’s preference for penuriousness over promiscuity in a formal confirmation theory. But it seems to me that, in any case, Glymour’s account is now both too tight and too loose. Furthermore, it seems to me that the type of modification Glymour would evidently like to make—adjusting the details of the account while retaining the classical approach to theories—is doomed to failure. In this paper, then, I would like to do four things: (in section I) point out briefly that the modified account is much more narrow-minded than Glymour suggests; (in section II) show that Glymour’s modifications have not succeeded in removing the account’s essential gullibility (and explain this failure by reference to bootstrapping’s reliance on the classical approach to theories); (in section III) argue that any similar modifications that did manage to solve the gullibility problem would induce unacceptable levels of narrow-mindedness; and (in section IV) discuss the possibility that some more drastically modified bootstrap-style account, purged of any connection to the classical approach to theories, would succeed in explaining evidential relevance. I will argue that, as of yet, we have seen no reason to believe that bootstrap-type relations have anything at all to do with evidential relevance.

I

Glymour discusses four of the examples given in (Christensen 1983). The first three were intended merely to show gullibility, and are indeed

¹In particular, the modified account never permits confirmation of a theoretical hypothesis having the form of a universally quantified conjunction.

disallowed by the formal condition Glymour proposes. The point of the fourth example, however, was somewhat more complex. It consisted of two sets of axioms which (to my mind, at least) bring with them quite different relevance-relations: a certain confirmation seems illegitimate on the first set of axioms, but quite legitimate on the second. However, the two axiom sets entail the same set of first-order consequences; on the classical approach, the two theories are identical. Since Glymour's account was based on the classical conception, I suggested that bootstrapping could not correctly assess evidential relevance in both cases, even if modifications were made to deal with my first three examples.

The two axiom sets had the following forms:

$$\begin{array}{ll} \text{I:} & \text{A1: } (x)(Fx \supset Gx) & \text{II:} & \text{A1: } (x)(Fx \supset Gx) \\ & \text{A2: } (x)(Fx \supset Hx) & & \text{A3: } (x)(Fx \supset (Gx \equiv Hx)) \end{array}$$

The problematic confirmation allowed evidence of the form

$$E: Fa \ \& \ Ha$$

to confirm A1. The original bootstrap account allowed confirmation in both cases; the modified account allows it in neither. Since this is not one of the cases in which Glymour says that his account is narrow minded, I assume that he does not share my intuition that the confirmation is legitimate for the second axiom set. Part of the reason for this, I suspect, is that the original interpretation I gave for the axioms was not maximally convincing on this score. What I would like to do here, then, is to give some examples of cases which show more clearly that evidence of the form *E* can indeed legitimately support hypotheses of the form A1, relative to a theory whose natural first-order axiomatization has the form II.

Consider a theory which consists of the hypothesis that all AIDS victims are infected by a certain virus *V*, along with the hypothesis that an AIDS victim will have antibody *B* in his blood just in case he is infected by virus *V*. The natural first-order representations of the hypotheses are as follows:

$$\begin{array}{ll} \text{H1:} & (x)(Ax \supset Vx) \\ \text{H2:} & (x)(Ax \supset (Bx \equiv Vx)). \end{array}$$

Now how would a scientist who believed both of these hypotheses go about confirming H1? It seems to me that she would very likely test AIDS victims for the presence of the antibody, and that she would (justifiably) regard finding the antibody in an AIDS victim as evidence for the hypothesis that all AIDS victims were infected with the virus.² However,

²This example is obviously simplified: for example, we might not think that every single AIDS victim would have the relevant antibody in his blood, especially considering the

on Glymour's new account, such a confirmation would be ruled illegitimate. Finding the antibody in AIDS victims would be ruled *irrelevant* to H1, even though our theory contains H2!

Similarly, suppose that a paleontologist believes that all halmasauruses were avid jumpers. How would he test this belief? Well, if he also believes that halmasauruses had fractured heel bones just in case they were avid jumpers, it seems to me that he might well regard halmasaurus fossils with fractured heel bones as confirming his hypothesis about their habits. The structure of the theory is identical to the previous one; again, Glymour's revised account would deny the legitimacy of the scientist's reasoning.

If my intuitions about confirmation are correct in the above cases, the narrow-mindedness induced by Glymour's modification of bootstrapping is indeed much more serious than he indicates.³ And it should also be kept in mind that, on the classical approach to theories, these examples are indistinguishable from the absurd confirmations Glymour modified his account to eliminate. Thus any further modification which resulted in allowing confirmations of the above form would (given Glymour's basic approach to theories) necessitate accepting clearly unacceptable confirmations along with them.

Still, if one thought that it was better for a formal confirmation theory to condemn the innocent than to let the guilty go free, one might think that rejecting a few legitimate confirmations—even quite a few—was an acceptable price to pay for a restriction that enabled the bootstrap account to provide a firm *sufficient* condition for evidential relevance.⁴ In the next section, however, I will argue that the revised account cannot even do this.

debilitating effect the disease has on the immune system. However, some of the reasoning employed in actual studies on AIDS is remarkably close to that described in the text. The following is from a recent book on AIDS: "[I]t was seroepidemiologic evidence that argued most strongly that HTLV-III does, in fact, cause the disease. For these studies, sera from a large number of patients with AIDS or ARC . . . and from normal control subjects were analyzed for antibodies to HTLV-III. . . . The results of testing over 1000 sera . . . demonstrated that over 90% of all AIDS and ARC patients were seropositive for HTLV-III. This was in contrast to the control group. . . ." (Gallo et al. 1985). Thus it seems to me that the suggestion that finding the antibody in AIDS victims provides *no* relevant evidence for H1 relative to a theory containing H1 and H2 is a clear distortion of scientific practice.

³In arguing for his own restriction on bootstrapping, Jan Zytow (1986) presents different, more complex scientific examples intended to show that Glymour's restriction eliminates desirable confirmations. (For more on Zytow's proposed condition, see the Appendix.)

⁴It should be noted, however, that narrow-mindedness (that is, failure to provide a necessary condition for confirmation) presents a serious obstacle to one of Glymour's main philosophical objectives. As Paul Horwich points out (Horwich 1983, pp. 55–56), failure to satisfy a merely sufficient condition for evidential relevance cannot provide an explanation of evidential irrelevance, something that Glymour clearly wants his account to do.

II

Before discussing my examples of gullibility, I would like to take a quick look at the restriction which is at the heart of the new, tougher, bootstrap condition. Officially, it reads as follows:

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*. (Glymour 1983a, p. 627)

Glymour explains that “*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” (1983a, p. 627). He says that this requirement is intuitively sensible, but he doesn’t say why.

Now it seems to me that there is a certain intuitive plausibility to *R*, at least on the surface, for the following reason: the computations it eliminates look *circular* in a certain way. An auxiliary hypothesis that violates *R* seems *dependent* on the hypothesis being tested, because in a sense it says nothing about one of its quantities that is not already said by the hypothesis being tested. Computations that actually use the tested hypothesis as an auxiliary (which Glymour originally intended to allow, giving the bootstrap theory its name) are merely a special case of this kind of circularity or dependence; and indeed, they will typically be disallowed in the revised bootstrap account.

Furthermore, it seems to me that it is precisely this intuitive dependence of the auxiliaries on the tested hypotheses that is at the root of the unacceptability of the confirmations Glymour is trying to disallow. Thus, if a syntactic test could be found for this intuitive dependence, the gullibility problem might be solved.⁵ But while the motivations behind the new condition are reasonable enough, it seems to me that Glymour’s test fails to disclose intuitive dependence in general, and thus that *R* fails to remove the essential gullibility of the bootstrap condition.

The first example of gullibility I want to discuss is simply a minor variation of one of the original examples *R* was introduced to eliminate. One of these original examples of gullibility consisted of a theory having two (intuitively) independent hypotheses of the following forms:

H1: $(x)(Rx \supset Bx)$

H2: $(x)(Fx \supset Gx)$

⁵The examples in the last section suggest, of course, that a condition that picked out the objectionable cases would also catch some legitimate cases in the same syntactic net. But while this certainly suggests that the intuitive dependence will prove elusive, it does not show that no interesting syntactic sufficient condition for acceptable auxiliaries can be found.

H1 was intended to represent the famous Raven Hypothesis; H2 was uninterpreted, but here we may take it to represent the hypothesis that only Gods can fly. The original bootstrap account permitted us, absurdly, to confirm H2 by spotting a flying black raven. The relevant computation made use of an auxiliary hypothesis H* that, intuitively speaking, is included in the theory only because H2 is:

$$H^*: (x)((Rx \supset Bx) \equiv (Fx \supset Gx)).$$

As Glymour points out, condition *R* successfully prevents us from using H* to confirm H2, because H2 entails the equivalence of H* with H1, which uses only a proper part of the vocabulary in H*.

Consider, however, a slight variation on our example. Suppose that our theory is supplemented by another hypothesis, say, that all winged things can fly:

$$H3: (x)(Wx \supset Fx).$$

The modified bootstrap account now allows us to confirm the hypothesis that only Gods can fly, simply by finding a black raven with wings! The computation goes as follows (the possible counterevidence would be a non-black, winged raven):

$$\begin{array}{ccccccc}
 & Fx & & Gx & & & \\
 H3 \leftrightarrow & \uparrow & \nearrow & \uparrow & \nwarrow & \leftrightarrow & (x)((Rx \supset Bx) \equiv (Wx \supset Gx)) \\
 & Wx & Rx & Bx & Wx & &
 \end{array}$$

This confirmation is certainly no less absurd than the original one. The strategy is in fact exactly the same as in the previous example; only the *direct* measurement of *F* has been replaced by *indirect* measurement through H3. Intuitively, the quantity *G* is still “measured” by assuming the hypothesis being tested. And the “possibility of disconfirmation” still rests on the possible failure of H1, an unrelated part of the theory, not on the possible failure of H2. But while the auxiliary hypothesis used in the right hand computation is *intuitively* dependent on H2, this dependence is not disclosed by the syntactic test R; thus the computation passes muster in the revised bootstrap account.⁶

The second example of gullibility that I would like to discuss is related

⁶To be precise, the auxiliary hypotheses used in the computations would have to be weakened to meet condition 6 in the revised account, which requires using the weakest possible hypotheses that satisfy the other five conditions. But since the condition requires weakening only to the extent that confirmation still occurs, my point is unaffected. Since fulfilling requirement 6 seems to me to detract significantly from the perspicuity of the examples, I’ll ignore it in the rest of this paper.

to an example of Aron Edidin's (1981). Suppose we have an equational theory consisting of the following hypotheses:

$$\text{H1: } x = y$$

$$\text{H2: } A = x$$

$$\text{H3: } B = x$$

A and B are to be thought of as observational quantities, x and y as theoretical; our theory gives us two ways of measuring x , and posits a relationship between x and y . As Edidin points out, we would certainly not want to count the following computation as confirming H1 by measuring A and B :

$$\begin{array}{ccc} & x & y \\ & \uparrow & \uparrow \leftrightarrow x = y \\ A = x \leftrightarrow & & x \\ & & \uparrow \leftrightarrow B = x \\ & A & B \end{array}$$

Edidin used the example—and variants of it—to argue for a restriction on bootstrapping to the effect that the auxiliaries used in a computation not be allowed to include, or even entail, the hypothesis being tested. (Clearly, Edidin's intuition here was to eliminate the kind of intuitive circularity we have been discussing.) Glymour, in *Theory and Evidence*, agreed with Edidin's suggestion that the computation was illegitimate, but not with Edidin's solution. Instead, he introduced a condition requiring that the possible counterevidence (in this example, any case where $A \neq B$) be consistent with the used auxiliaries (see Glymour 1980, p. 117 fn). On either proposal, the above computation is disallowed.

Now the hypotheses used in the above computation are intuitive axioms of the theory, and one of them is the very hypothesis being tested. But computations need not wear their strategies on their sleeves. The following compressed computation is in a clear sense equivalent to the first, but it does not *explicitly* use H1, eliminating the middle step by transitivity:

$$\begin{array}{ccc} & x & y \\ A = x \leftrightarrow & \uparrow & \uparrow \leftrightarrow B = y \\ & A & B \end{array}$$

This streamlined version of the computation meets Glymour's *T & E* requirements, and also meets the requirement Edidin proposes in response to the expanded version. And—importantly for our present purposes—it also meets the more stringent requirements of Glymour's revised bootstrap account. Nevertheless, it seems just as clear a case of gullibility as does Edidin's original example. The same evidence is used in the same way to confirm the same hypothesis relative to the same theory. The

“measurement” of y still depends intuitively on assuming the truth of the tested hypothesis; although H_1 is not mentioned explicitly, the auxiliary used in the right hand computation is intuitively dependent on it. The “possible counterevidence” would still have to result from a breakdown of H_2 or H_3 , not from failure of the hypothesis being tested. In short, the gullibility problem pointed out in the original example remains unabated.

To bring the gullibility problem into even sharper focus, let us consider a new interpreted example, this one set in the first-order version of the bootstrap account, rather than the equational version. Suppose that I believe that all and only Zoroastrians will have eternal life:

$$H_1: (x)(Ex \equiv Zx).$$

Also, suppose that I have a couple of different ways of identifying Zoroastrians. I believe that anyone wearing a sudra is a Zoroastrian

$$H_2: (x)(Sx \supset Zx);$$

and that all and only Zoroastrians pray to Ahura-Mazda

$$H_3: (x)(Px \equiv Zx).$$

Now suppose that I see my next-door neighbor wearing a sudra while praying to Ahura-Mazda. Should I regard this as confirming my belief that eternal life is the reward of all and only Zoroastrians?

According to the revised bootstrap account, such “confirmation” is perfectly proper.⁷ Yet to my mind, this is as clear a case of evidential irrelevance as any cited by Glymour. Again, one of the hypotheses used in the confirming computation is intuitively dependent on the tested hypothesis; and again, this is not disclosed by the syntactic test built into the revised bootstrap account.

The impression that confirmation is improperly circular here can perhaps be brought out a bit more sharply by considering the following variant on the case: The man two doors down, being somewhat less favorably disposed toward Zoroastrianism than I am, might well share my beliefs about how Zoroastrians can be identified (H_2 and H_3), while rejecting my H_1 in favor of the dramatically opposite hypothesis that all and only Zoroastrians will *fail* to have eternal life:

$$H_1^*: (x)(Zx \equiv \sim Ex).$$

⁷It would look like this (where the confirming evidence is Sa & Pa , and the possible counterevidence is Sa & $\sim Pa$):

$$\begin{array}{ccc} Zx & Ex & \\ H_2 \leftrightarrow \uparrow & \uparrow & \leftrightarrow (x)(Px \equiv Ex) \\ Sx & Px & \end{array}$$

On Glymour's new account, this man can claim that the very same evidence I cited (our neighbor's wearing a sudra while praying to Ahura-Mazda) supports *his* views on the long-term prospects for Zoroastrians!

It might be objected that while the bootstrap account is clearly too gullible here, this is hardly a realistic *scientific* example; and that true scientific examples with this structure would be so rare that this gullibility should be seen as unfortunate, but acceptable. However, I see no reason to doubt that "confirmations" of this type are often possible in real scientific theories (though no scientist would bother to consider them, of course). Laws which assert a determinate enough relation between two theoretical quantities will often be "confirmable" simply by finding two different ways of measuring one of the quantities. In fact, counterintuitive confirmations of a very similar type will even be possible in the case of Kepler's laws, which is Glymour's central example of a scientific case of evidential irrelevance.

Kepler's third law states that some constant ratio K will hold between the period and the $3/2$ power of the mean distance from the sun of any two planets. Clearly, observations of a single planet should not confirm the third law. In *Theory and Evidence*, Glymour claimed that even if we treated K as a single quantity, we wouldn't be able to bootstrap-confirm the third law by observations of a single planet (say, Mars). If we tried, for instance, to use observations of Mars to determine not only $K(\text{Mars})$ but also $K(\text{Venus})$ (in order to test the third law), we would have to rely on the third law itself; and this circularity would necessitate violation of one of the conditions of bootstrapping.

As it turned out, computations involving the objectionable kind of circularity did not always violate the original version of the bootstrap account (see Christensen 1983); indeed, Glymour's new condition is in part intended to block these illegitimate "confirmations" of Kepler's third law (see Glymour 1983a). However, as we will see, the revised version of bootstrapping is vulnerable to essentially the same type of example. I would like to show this in a way that will first demonstrate how Glymour's restriction is supposed to work, in the hope that this will make clearer the reason for its eventual failure.

Let us begin by supposing merely that Kepler had more than one way of measuring the period and mean distance from the sun, and hence the K -value, of a planet. I'll represent these methods of measurement by H_1 and H_2 , and Kepler's third law by H_3 :

- H_1 : $(x)(O_1x = Kx)$
 H_2 : $(x)(O_2x = Kx)$
 H_3 : $(x)(y)(Kx = Ky)$.

A first stab at constructing a counterintuitive confirmation might look like this:

$$H_1 \leftrightarrow \begin{array}{cc} Kx & Ky \\ \uparrow & \uparrow \\ O_1x & O_2x \end{array} \leftrightarrow (x)(y)(O_2x = Ky)$$

Our data would consist of two observations of the same planet—say, Mars. One observation would determine the value of $O_1(\text{Mars})$, the other the value of $O_2(\text{Mars})$. We would use the left hand side of the above computation to compute $K(\text{Mars})$, in the natural way; but with the right hand side of the computation, we would calculate, say, $K(\text{Venus})$. This would, of course, give us a positive instance of the third law when the two measurements of Mars agreed with one another, and a negative instance when the measurements of Mars differed.

Now there are two difficulties with this initial attempt at generating a counterexample to Glymour's account. First, it violates the requirement originally laid down in *Theory and Evidence* (though apparently dropped by Glymour in the revised account) that the possible counterevidence be consistent with the used auxiliaries. Second, it violates the new restriction R , since H_3 entails that the hypothesis used in the right hand computation is equivalent to H_2 , whose vocabulary includes one less quantified variable.⁸

The second of these problems in particular does seem to bear out Glymour's reasons for instituting the new requirement. The strange hypothesis used in the right hand computation depends on the tested hypothesis in an intuitively obvious way, and Glymour's vocabulary restriction appears to pick up on just this fact: one might put it, loosely, by saying that the strange hypothesis says nothing about the extra quantified variable that is not said by the tested hypothesis. In this way, the syntactic structure of the strange hypothesis does look like an indicator of what we intuitively express by saying that part of the content of the strange hypothesis "comes from" or is "dependent on" the tested hypothesis. And it is this intuition that is at the root of our rejecting the computation as unacceptably circular.

Unfortunately, intuitive dependence between hypotheses is hard to catch syntactically. In this case, it turns out that a very weak assumption about our astronomical theory allows the construction of an intuitively unacceptable confirmation that does not violate the new vocabulary restriction

⁸I am supposing here (with no good reason) that Glymour intends quantified variables to be included in a sentence's "essential vocabulary". If not, the case just discussed already constitutes a counterexample to the revised account.

(or the *Theory and Evidence* consistency-of-counterevidence condition, or Eddin's condition either). We need only assume that the astronomical theory contains some observation-predicate P (say, "is bright") that is true of Venus. We then consider the following computation:

$$\begin{array}{ccc}
 & K(x) & K(y) \\
 H_1 \leftrightarrow & \uparrow & \uparrow \leftrightarrow (x)(y)(n)((O_2x = n \ \& \ Py) \supset Ky = n) \\
 & O_1(x) & O_2(x)
 \end{array}$$

An evidence sentence stating $P(\text{Venus})$, and giving equal values for $O_1(\text{Mars})$ and $O_2(\text{Mars})$, will yield (via the above computation) equal values for $K(\text{Mars})$ and $K(\text{Venus})$, bootstrap-confirming the third law. If the values for $O_1(\text{Mars})$ and $O_2(\text{Mars})$ differ, the computation will yield a negative instance of the third law.

Nevertheless, it is perfectly clear that the third law cannot really have been tested at all by this computation. No observation of Venus has been made which bears even remotely on the third law holding; the "measurement" of $K(\text{Venus})$ is intuitively parasitic on the *assumption* that the third law holds. Any "confirmation" obtained through this computation will thus be trivially circular. The "possibility of disconfirmation" is also fully independent from any failure of the third law; essentially, it rests on the possible failure of an unrelated part of the theory (the part which says that the two measurements of Mars will agree with one another).

It seems to me that this example clearly shows that the gullibility of the modified account does indeed infect the kinds of cases Glymour is centrally interested in explaining. It also shows how the restriction R , however well motivated, fails to accomplish its purpose. As in the simpler and more artificial cases, R is just not capable of capturing the intuitive dependence between those hypotheses used in a computation, and the hypothesis being tested. Thus it seems that the basic source of gullibility in the original account remains; and, as a result, the added restriction completely fails to solve the bootstrap account's problem.⁹

⁹It should be noted that attempts have been made to solve the gullibility problem using restrictions other than R . Such restrictions have been put forth by Jan M. Zytow (1986) and by John Earman and Glymour (1988). Neither of these conditions succeeds in disallowing counterintuitive confirmations of the sort discussed above. For details, see the Appendix.

I should also note here that Thomas R. Grimes (1987) has pointed out that a very different sort of "promiscuity" infects bootstrapping. Grimes shows that almost any piece of evidence will confirm almost any theory, *relative to some true theory or other*. The "theories" Grimes invokes to make his point are essentially conditionals formed from the evidence sentences and instances of the tested hypothesis—not the kind of theory relative to which we typically want to assess confirmation. Nevertheless, Grimes' point raises interesting questions, especially as bootstrapping gives us no well-developed way of assessing the confirmation of the theories relative to which we confirm individual hypotheses.

III

The examples discussed so far show that satisfaction of the modified bootstrap account is neither necessary nor sufficient for confirmation. One could, of course, always modify the account still further; and it might be hoped that some such further modification could really eliminate (or markedly reduce) the account's gullibility, providing a sufficient condition for evidential relevance. However, some of the examples discussed in the last section seem to me to provide good reason for pessimism on this score.

The examples in section I showed that, given Glymour's basic approach to theories, his exclusion of certain unacceptable confirmations was necessarily accompanied by exclusion of other, perfectly reasonable, confirmations. That point having been made, we went on in section II to see whether rejecting the reasonable confirmations had at least allowed the modified account to provide a usable *sufficient* condition for evidential relevance; we found that, in fact, it had not. But nothing in section II militated against the possibility that some re-modified bootstrap account could succeed in doing this.

Before discussing this question, I would like to make a simple methodological observation. It is obvious that we could give a *trivial* bootstrap-style sufficient condition for evidential relevance—a condition that excluded all confirmations would do that. The interesting question is whether a bootstrap-style account can provide an *interesting* or *useful* sufficient condition: a condition that excludes the undesirables while also letting in a significant fraction of the confirmations we favor. An interesting bootstrap account would have to be able, quite frequently, to *discriminate* between acceptable and unacceptable confirmations.

With this in mind, it is clear that the examples discussed in section I already give us some reason for pessimism about the prospects for a re-revised Glymourian account. They constitute cases in which no account which shares Glymour's basic approach to theories can discriminate between reasonable and unacceptable confirmations. In the present section, I want to show how some examples discussed in section II present additional, parallel difficulties for Glymour's approach.

Consider the "streamlined" variation on Edidin's example, discussed on page 650. If, instead of the axioms given there, the natural axioms of our theory were as follows:

- H₁: $x = y$
- H₂: $A = x$
- H₃: $B = y$;

then testing the first hypothesis by measuring A and B would be entirely appropriate—in fact, it would seem to be a virtual paradigm of the sort of confirmation Glymour envisions. Yet the set of consequences entailed by these axioms is identical to the set entailed by the axioms discussed above, for which the same confirmation was clearly—again, almost paradigmatically—illegitimate. No account that takes into account only the logical structure of the set of first-order consequences of a theory will be able to discriminate between these virtual caricatures of evidential relevance and irrelevance.

Similarly, consider the following anthropological theory: Languages in group Z were spoken by all and only tribes in group E :

$$H_1: (x)(Zx \equiv Ex);$$

Any tribe using script S spoke language Z :

$$H_2: (x)(Sx \supset Zx); \text{ and}$$

E -tribes lived in structures of type P , and they were the only tribes to do so:

$$H_3: (x)(Px \equiv Ex).$$

Suppose that we hold this theory, and that we find remains of a tribe who lived in P -structures and also used script S . It seems that we should be able to use this evidence in a perfectly reasonable way to confirm H_1 . In fact, our reasoning will look strikingly bootstrappy: we will use H_2 to infer that our tribe spoke language Z , and use H_3 to infer that our tribe was of group E ; we will then have a positive Hempelian instance of our tested hypothesis. (It might also be noted that we could have found remains of a tribe who used script S without living in P -structures, thereby disconfirming H_1).

The point of all this is, of course, that this perfectly reasonable confirmation is identical, from the point of view of logical structure, to the absurd confirmation in the “Zoroastrianism” example discussed above on page 651. Again, any way of tightening up the bootstrap account to eliminate the latter will eliminate the former at the same time. No account of confirmation that analyzes evidential relevance in terms of the logical structure of the set of first-order consequences of a theory can possibly *discriminate* between such cases of clear relevance and clear irrelevance. In light of these examples, as well as those discussed in section I, it seems highly unlikely that bootstrapping, with its classical approach to theories, can furnish us with a sufficient condition for evidential relevance that is at all useful or interesting.¹⁰

¹⁰I would like to make it clear that to demand that a satisfactory account of confirmation

IV

The considerations in the last section suggest that the prospects for a further-revised version of bootstrapping are not promising. However, my reasons for pessimism on this score have been predicated explicitly on the assumption that bootstrapping remain tied to the classical approach to theories. Nothing in the above examples throws doubt on the prospects for a further-revised bootstrap-style account that, say, took “natural axiomatizations” of theories into account. Such an account would not be committed to the equivalence of the pairs of theories that I have argued must be treated non-equivalently.

Of course, such an account would have to be compared, for instance, to a version of hypothetico-deductivism that was also allowed access to the richer conception of theories. This is important, for if hypothetico-deductivism could account for evidential relevance, bootstrapping would lose much of its appeal. It is worth remembering that Glymour’s main claim for the superiority of his account over the hypothetico-deductivist’s is that the latter cannot account for relevance; yet he admits (1980, p. 39) that the hypothetico-deductivist might well be able to account for relevance if allowed to pick certain axiomatizations as privileged. Thus admitting privileged axiomatizations, while it may well be a step in the right direction for confirmation theory, might still not be a step which rendered bootstrapping particularly interesting.

Another alternative to the classical approach to theories would replace classical first-order logic with some alternative, such as a relevance logic or a logic of counterfactuals. Again, such an approach to theories might well allow differential treatment of the theory-pairs discussed above. And in fact, it has recently been argued (Waters 1987) that relevance logic can save hypothetico-deductivism from the irrelevance problems pointed out by Glymour. If the hypothetico-deductivist’s problems with relevance are similar to the bootstrapper’s (as I have suggested), then if Waters’ account furnishes hope for hypothetico-deductivism, there may be hope

make the discriminations discussed above does not amount to denying the “equivalence condition”, which entails that confirmation relative to a theory T should not be affected by substituting a logically equivalent theory T' for T (see Earman and Glymour 1988, p. 262).

The examples discussed above involve pairs of theories whose natural first-order axiomatizations yield identical sets of consequences. I do claim that we cannot substitute these *theories* for one another without affecting confirmation. But this requires abandoning the equivalence condition only if we adopt the classical approach of identifying theories with the set of consequences of their natural first-order axiomatizations. There are, of course, alternatives to this approach, such as using a different logic to represent our theories (this option—which actually constitutes one possibility for saving some form of bootstrapping—will be briefly discussed below).

for bootstrapping along the same road.¹¹ Again, however, it is far from clear that such a readjustment will leave bootstrapping in any better position relative to its main competitor.

Thus at this point we do not know whether we could obtain useful results from a bootstrap-style account that jettisoned the classical approach to theories. But while we have seen no argument that such results are impossible, we have as yet seen no evidence suggesting that the approach is likely to produce a bootstrap account which could account for relevance better than hypothetico-deductivism. Thus I must remain guardedly pessimistic about the prospects for revising bootstrapping along these lines.

Before concluding, however, I would like to mention one more move that might be made on bootstrapping's behalf, a move which does not depend on abandoning Glymour's approach to theories. It has been suggested (Edidin 1981) that Glymour never claimed to be giving a *complete* account of confirmation. Perhaps bootstrapping will explain relevance only in combination with some other factors, or only for certain restricted classes of theory. For example, it might be held that bootstrapping does correctly determine relevance relations for deductively closed sets of first-order sentences; and that our intuitions to the contrary in the cases discussed above relate to the theories which I stated in English and which aren't adequately represented in the obvious first-order way.

The first thing to be noticed about this last way of seeing the bootstrap account is that, given the points made above about the Kepler's laws example, we would have to limit the strategy so severely as to prevent it from applying to the central case Glymour adduces in its defense. Such a limitation should, to say the least, raise serious questions about the interest of the account.

This brings up a more general point that is important here. It is true that, to the extent the claims made for bootstrapping are of a more limited nature, it will be harder to show the condition defective by means of a few simple examples. But care must be taken not to insulate the account from failure by depriving it of determinate content. We cannot simply assume that when the account accords with our intuitive judgments, it explains them; while when it disagrees with us, some other factor is operating in addition to—or instead of—bootstrapping. If a limited bootstrap account is to have any substance, it must include a general account of its own limitations, so that we may apply the condition independently of our intuitions about relevance in particular cases.

¹¹In fact, Glymour reports that “[Kevin] Kelly has shown that Christensen's difficulty does not arise when [bootstrapping] is developed in the context of relevance logic” (1983b, p. 5).

Now I have no argument showing such an account to be impossible. Perhaps there are other factors which help explain relevance, yet which are not so explanatorily powerful as to leave the bootstrap strategy without significant work to do. At this point in the discussion, however, no such account has even been hinted at; and until one is, I see little reason to think that such an account will be found. So, again, I feel somewhat pessimistic about the prospects for this sort of revision of bootstrapping.¹²

I would like to conclude by saying something about how we should go about assessing future incarnations of the bootstrap account. One of the main attractions of the account has been its ability to represent real scientific cases of confirmation by relevant evidence. But it is important to realize that in order to argue that some future version of bootstrapping is responsible for our judgments of evidential relevance, it will not be sufficient to point to a few—even quite a few—prominent scientific arguments that are roughly representable in bootstrap style. It will have to be shown that we generally refrain from performing or taking seriously tests that violate the account. And it will have to be shown that we (at least tend to) perform and take seriously tests that satisfy the account, when they are easily available.

This last stricture applies especially to any tests that use auxiliaries which are intuitively included in a theory because the tested hypothesis is included. As is clear from the examples discussed above, such “tests” are not hard to think up; thus, to the extent that any such tests are legitimate, they will surely have been performed and taken seriously.

I suspect, however, that the extent to which such confirmations are legitimate is very small, and thus that any viable account of evidential relevance will have to find a way of identifying and rejecting this sort of test. Auxiliaries which depend on the hypothesis being tested are the source of the trivializing circularity that characterizes all of the above-discussed examples of gullibility. Such auxiliaries are at the root of the problem that both the bootstrap account and the hypothetico-deductive account have with relevance; and they have so far proved resistant to characterization by standard first-order methods. In assessing the adequacy of any future version of bootstrapping, we must clearly take very special care to see whether it handles this particular sort of case correctly.

At present, we have seen no version of bootstrapping that comes close

¹²A new proposal along somewhat similar lines has recently been made by Aron Edidin (1988). He suggests that we understand the purpose of bootstrapping in a new, and much more modest, way. On Edidin's proffered interpretation, we should not ask bootstrapping to make the kind of discriminations discussed above; thus Glymour should never have worried about the gullibility examples in the first place! Now it seems to me that the suggested weakening of the claims made for bootstrapping would rob the account of most, if not all, of its interest. However, detailed discussion of Edidin's proposal must await another occasion.

to accomplishing this. Although there are, to be sure, many examples of correct scientific argument that are representable in roughly the bootstrap way, the same can be said for classical hypothetico-deductivism. In fact—and this strikes me as very important—we have seen *no* significant class of theories, real or artificial, in which bootstrapping can reliably discriminate between relevant and irrelevant confirmations. Thus at present, we have seen little reason to believe that bootstrapping will play any useful role at all in explaining evidential relevance.¹³

APPENDIX

Some Other Attempts to Solve the Gullibility Problem

In “What Revisions does Bootstrap Testing Need?”, Jan M. Żytkow (1986) suggests that the counterintuitive confirmations described in Christensen (1983) be eliminated by the following requirement on computations:

Evidence E_1 is logically possible such that by use of the same computational procedure:

- (1) E_1 disconfirms H , and
- (2) for every subtheory S in T such that S does not contain any consequence of H (tautologies excluded), E_1 does not disconfirm S . (Żytkow 1986, p. 105)

Żytkow’s intuition is that E_1 should not disconfirm any part of T that is *independent* of H ; independence is supposed to be captured by the “no shared consequences” clause in

¹³It might seem that one exception to my last claim could be provided by Glymour’s bootstrap-style analysis (1983b) of fault-detection in logic circuits. A logic circuit can be represented by a system of equations specifying the output of each gate as a function of its inputs. This system of equations is in effect a theory of the circuit. Failure of a particular gate to conform to the equations is thus analogous to falsity of a particular hypothesis in a theory; and the problem of locating a particular faulty gate by measuring inputs and outputs of the circuit is analogous to selective confirmation or disconfirmation of particular hypotheses in a theory. Glymour describes a bootstrap-style procedure for solving this sort of problem, and notes that this procedure is equivalent to the Boolean derivative tests that are used in practice to locate gate-failures. Thus it would seem that in one real (if small) set of cases, a bootstrap-type procedure is useful in determining confirmational relevance relations.

However, a closer look at the actual bootstrap procedure employed by Glymour casts doubt on even this example. It turns out that Glymour’s procedure involves not only bootstrap testing relative to the theory determined by the entire circuit-diagram in question, but another bootstrap test as well. This second bootstrap test is relativized to a special *subtheory* of the original one—a subtheory that is determined by deleting certain sections from the circuit diagram, then taking the set of equations determined by the truncated diagram. Now the individual sections of a circuit diagram correspond intuitively to the natural axioms of the theory determined by the diagram. Thus the special subtheory used by Glymour is essentially a subtheory determined by removing some of the natural axioms of the original theory. In other words, the special bootstrap-style procedure Glymour is employing in this set of examples is dependent on taking a certain natural axiomatization of the original theory as privileged. This is, of course, one way of dealing with bootstrapping’s relevance problem; but, as noted above, once we make bootstrapping dependent on assuming privileged axiomatizations, the purported advantage of bootstrapping over H-D accounts may well evaporate. In fact, in this particular case Glymour explicitly notes that the results of the bootstrap procedure are *equivalent* to those obtainable by a H-D procedure which is allowed to take certain axioms as privileged!

(2). He illustrates the workings of his condition in dealing with an example involving a theory containing both the Raven Hypothesis and a simple formulation of pantheism:

$H_1: (x)(Rx \supset Bx)$

$H_2: (x)Gx$

On Glymour's original account, a black raven confirms pantheism, and a non-black raven disconfirms it. Żytkow writes:

In [this] example, H_1 does not contain any consequences of H_2 and should not be disconfirmed by E_1 . However, possible counterevidence for H_2 , that is $Rb \ \& \ \neg Bb$, disconfirms H_1 . Therefore, E_1 is not admissible by [clause (2) of the above condition], thus [the example] is no longer an example of bootstrap confirmation. (Żytkow 1986, p. 105)

Unfortunately, the argument rests on the false assumption that H_1 contains no consequences (that is, shares no logical consequences) of H_2 . H_1 and H_2 share such non-tautologous consequences as their disjunction and $(x)((Rx \supset Bx) \vee Gx)$. As Żytkow has acknowledged (in correspondence), these shared consequences render the condition vacuously satisfied, in this case and in the others it was designed to exclude. Thus it seems that this condition does not succeed in eliminating (or even reducing) the gullibility of the bootstrap account.

A more recent attempt at curbing promiscuity is made by John Earman and Glymour (1988), in a reply to Żytkow's paper.¹⁴ Earman and Glymour propose imposing the following restriction on computations (here i indexes the computation by which quantity Q_i is computed by means of auxiliary hypothesis T_i ; H is the confirmed hypothesis; E is the confirming evidence; and E' the possible disconfirming evidence):

(R') H does not entail that for some i , $\vdash T_i \leftrightarrow T'_i$ where E and T'_i do not jointly entail a value for Q_i and where E' and T'_i are inconsistent. (Earman and Glymour 1988, p. 263)

Earman and Glymour acknowledge that R' is somewhat undermotivated. But like the original condition R, it does succeed in disallowing the problematic computations described in Christensen (1983). However, it also resembles R in failing to exclude the problematic computations described in section II above. Thus it, too, fails to solve bootstrapping's essential problem with gullibility.

REFERENCES

- Christensen, D. (1983), "Glymour on Evidential Relevance", *Philosophy of Science* 50: 471–481.
- Earman, J., and Glymour, C. (1988), "What Revisions does Bootstrap Testing Need? A Reply", *Philosophy of Science* 55: 260–264.
- Edidin, A. (1981), "Glymour on Confirmation", *Philosophy of Science* 48: 292–307.
- . (1988), "From Relative Confirmation to Real Confirmation", *Philosophy of Science* 55: 265–271.
- Gallo, R. C., Shaw, G. M., and Markham, P. D. (1985), "The Etiology of AIDS", in V. T. DeVita, Jr., S. Hellman and S. A. Rosenberg (eds.), *AIDS: Etiology, Diagnosis, Treatment, and Prevention*. Philadelphia: Lipincott.
- Glymour, C. (1980), *Theory and Evidence*. Princeton, New Jersey: Princeton University Press.

¹⁴Earman and Glymour also criticize Żytkow's condition on the grounds that it is "nearly vacuous", because the "no shared consequences" requirement will be satisfied only in trivial cases. However, their argument to show this rests on the mistaken claim that S and H will always share non-tautologous consequences, unless either S and H are inconsistent, or one of them is tautologous itself. To see that this claim is incorrect, take S to be $(P \supset Q)$ and H to be $(\neg P \supset R)$. Thus although Żytkow's condition is far too weak to solve the gullibility problem, it is not so nearly vacuous as Earman and Glymour suggest.

- . (1983a), “Revisions of Bootstrap Testing”, *Philosophy of Science* 50: 626–629.
- . (1983b), “On Testing and Evidence”, *Minnesota Studies in the Philosophy of Science* 10: 3–26.
- Grimes, T. R. (1987), “The Promiscuity of Bootstrapping”, *Philosophical Studies* 51: 101–107.
- Horwich, P. (1983), “Explanations of Irrelevance”, *Minnesota Studies in the Philosophy of Science* 10: 55–65.
- Waters, C. K. (1987), “Relevance Logic Brings Hope to Hypothetico-Deductivism”, *Philosophy of Science* 54: 453–464.
- Żytkow, J. M. (1986), “What Revisions does Bootstrap Testing Need?”, *Philosophy of Science* 53: 101–109.