RICHARD OTTE

CRITICAL REVIEW: BRIAN SKYRMS, *CAUSAL NECESSITY*

(Received 15 November, 1982)

In a recent book, *Causal Necessity*, Brian Skyrms has attempted to deal with many of the major problems facing philosophers today.[1] The central idea of his book is that invariance is the key to understanding many of these problems. The idea of invariance is applied to problems such as randomness, epistemic probabilities, confirmation, conditionals, and decision theory. In this article I will briefly present the essence of his position on invariance, and then critically analyze it.

The first part of Skyrms' book deals with propensities and statistical laws. Skyrms believes that propensities are the probabilities that play a role in statistical laws, and his discussion of statistical laws assumes this. Statistical laws tell us that certain systems have a stable probability; these stable probabilities are propensities. Skyrms discusses in detail what it means for a probability to be stable, and he defines a notion of resiliency which is supposed to capture the idea of invariance and stability. Resiliency is defined as:

Resiliency of $\Pr(q)$s being $\alpha = 1 - \mathrm{Max}_i|\alpha - \Pr_j(q)|$ over $p_1 \ldots p_n$ (where the $\Pr_j$s are gotten by conditionalizing on some truth-functional compound of the p_is which is logically consistent with both q and its negation) (pp. 11–12).

In this definition the p_is are properties or experimental factors which are considered relevant to the occurrence of q. Resiliency measures the independence of q and these factors; thus resiliency is a measure of stability, independence, and invariance. If the resiliency of a certain proposition is 1, then we know that the proposition is necessary or invariant. Degrees of resiliency less than one correspond to cases of approximate independence or approximate invariance. Thus we can look at the resiliency of a proposition to determine how close we are to the ideal.

There are, however, some problems that arise with this definition of resiliency. One problem concerns the way in which the degree of resiliency is measured. Resiliency is measured by the maximum difference between α and the probability of q conditional on various truth functional combina-

Philosophical Studies **44** (1983) 425–433. 0031–8116/83/0443–0425$00.90

tions of the p_is. But there are cases in which this difference is not an adequate measure of the degree of invariance of a proposition. Events with a very low probability can be quite variant, and yet be highly resilient according to Skyrms' definition. Suppose that we have a certain atom with a large half life; the probability of this atom decaying in a certain time interval, say one year, will be quite small. Even if conditional on one of the p_is the probability of decay is doubled, the decay of the atom is highly resilient, because of the small difference between the actual probability values. If $\Pr(q)$ is low enough, the difference between $\Pr(q)$ and $2\Pr(q)$ will be very small. A small difference between two numbers is compatible with a large ratio between them. This seems to indicate that the difference between these two numbers cannot be an adequate measure of resiliency or invariance.

Contrast the previous example with another example of an atom which has a much shorter half life. The probability that this atom may decay in the same time period may be much higher, let us say around 0.25. If, relative to one of the p_is, the probability of this atom decaying also doubles, it will not be a very resilient probability. In this case the resiliency would be 0.75, which is not nearly as high as the resiliency of the other example. The only difference between this example and the previous one is the actual probability values involved. It seems that if one of these is highly invariant, so should the other be. The same atomic theory is behind both of them: if one of these probabilities is a propensity, they should both be propensities.

Another example which makes the same point is as follows. The probability of an atom decaying in a certain time interval may be very low, and thus even if there is a p_i which doubles that probability, it will be highly resilient. However, the probability that the same atom will decay in a much larger time interval will be much larger, in which case it is likely that the decay of the atom will not be highly resilient. The only difference in this example is the longer time intervals involved, which results in different probability values. But it seems clear that the longer time interval does not affect the independence of the propositions involved; all it does is change the probability values. Skyrms' definition of resiliency is defective because it does not account for cases such as these.

The definition of resiliency that Skyrms first presented was only applicable to finite probability spaces. However, Skyrms believes that nonstandard measure theory enables us to talk meaningfully about infinite probability spaces. Accordingly, we must have a definition of resiliency that is applicable to these idealizations. One nice part of Skyrms' book is the short introduction that he gives to nonstandard measure theory. For our purposes, all we need to realize is that the range of the probability function will be the nonstandard reals, which are the reals with the addition of infinitesimals. Any nonstandard real number has a unique representation as the sum of a standard real number and an infinitesimal; thus Skyrms is able to refer to the standard part of a nonstandard real number. Given this, Skyrms defines resiliency for infinite probability spaces as follows:

The resiliency of q = greatest lower bound of the set of all numbers, n, such that n is the standard part of Pr(q given p) for some p consistent with q (p. 76).

We should notice that this definition is really a definition of the resiliency of Pr(q) being 1. Although Skyrms does not do so, we could also define the resiliency of Pr(q) being α as follows:

> The resiliency of Pr(q) being α = 1 − the least upper bound of the set of all numbers, n, such that n is the standard part of $|\alpha - \text{Pr}(q \text{ given } s)|$, for s consistent with q and not-q.

This more general definition is actually more useful than the one Skyrms presents, because we want to be able to talk about the resiliency of propositions expressing probabilities, and not just universal generalizations. Without this generalization, in infinite probability spaces we could only talk about the resiliency of Pr(q) being 1, which would be a severe limitation.

One interesting fact that follows from Skyrms' definition of resiliency, is how difficult it is for a proposition to be highly resilient. This seems to be a property that holds in general for rich languages. Skyrms says that a language, along with a probability distribution on it, is *spread out* if and only if every proposition of the language with positive probability is entailed by a proposition with infinitesimal probability (p. 77). We should notice that most rich languages will be spread out. To see this, consider the example of a dart that is thrown at the unit interval. For any subset of the unit interval which the dart has a positive probability of hitting, there is a point contained in that interval such that the probability of the dart hitting it is infinitesimal, and yet the dart hitting it entails that the dart landed in the interval. Skyrms also says that we say that a certain value is *almost equal* to a certain value if and only if it differs at most infinitesimally from that value. Using this terminology, Skyrms (p. 77) proves the following theorem:

If a language, together with a probability distribution on it, is spread-out, then for any proposition in that language its resiliency is equal to zero or its probability is almost equal to one.

The purpose of this theorem is to show how difficult high resiliency is to come by in very rich language. This is a desirable result, because it guarantees that if a proposition is resilient, then it is very stable or invariant.

Unfortunately, one cannot generalize the above theorem to claim that the resiliency of $\Pr(q)$ being α is equal to zero or its probability is almost equal to α. The reason for this is that the smallest possible value of the resiliency of $\Pr(q)$ being α depends on the value of α. Consider Skyrms' original definition of resiliency. If α is 0.5, then the smallest possible value the resiliency of $\Pr(q)$ being α can have is 0.5, because the maximum value of $\alpha - \Pr_j(q)$ is 0.5. In order for this generalization to be true, the definition of resiliency must be normalized in some way.

However, a related theorem can be proven. Before we prove it we should notice that the smallest possible value the resiliency of $\Pr(q)$ being α can have is equal to $0.5 - |0.5 - \alpha|$. Given this, we can prove the following theorem:

If a language, together with a probability distribution on it, is spread out, and neither α nor $\Pr(q)$ is almost equal to one or zero, then the resiliency of $\Pr(q)$ being α is the smallest possible value that it can have, *even if* $Pr(q) = \alpha$.

Proof: Suppose that $\alpha \geqslant 0.5$. Then $\Pr(\bar{q})$ is finite, since $\Pr(q)$ is not almost equal to one. Since the language is spread out, there is a q^* with infinitesimal probability such that q^* implies q. Then $q^* \vee \bar{q}$ is consistent with both q and not-q, and $\Pr(q \text{ given } q^* \vee \bar{q})$ is at most infinitesimal. So $|\alpha - \Pr(q \text{ given } q^* \vee \bar{q})|$ is almost equal to α, and thus the resiliency of $\Pr(q)$ being alpha equals $1 - \alpha$, which is equal to the smallest possible value that it can have. Now suppose that $\alpha < 0.5$. Then $\Pr(q)$ is finite, since it is not almost equal to 0. Since the language is spread out, there is a $\bar{q}^*$ with infinitesimal probability such that $\bar{q}^*$ implies $\bar{q}$. Then $q \vee \bar{q}^*$ is consistent with both q and not-q, and $\Pr(q \text{ given } q \vee \bar{q}^*)$ is almost equal to one. So $|\alpha - \Pr(q \text{ given } q \vee \bar{q}^*)|$ is almost equal to $1 - \alpha$. Thus the resiliency of $\Pr(q)$ being α equals α, which is the smallest possible value that it can have.

This theorem is important, because it shows that the definition of the resiliency of $\Pr(q)$ being α in infinite probability spaces is totally unsatisfactory. One reason we wanted a definition of resiliency was in order to deal with statistical hypotheses. But according to this theorem, every statistical hypothesis has the lowest degree of resiliency possible; even in the cases in which

$\Pr(q)$ equals α, its resiliency is as low as it can be. This result shows that Skyrms' treatment of resiliency in infinite probability spaces cannot easily be generalized to account for statistical laws. Thus in the following discussion we shall not deal with the resiliency of $\Pr(q)$ being α, where α is not equal to 1. We shall restrict our discussion to Skyrms' original definition of the resiliency of $\Pr(q)$ in infinite probability spaces.

However, given Skyrms' definition of resiliency in infinite probability spaces, a problem arises when we consider whether it is possible for any contingent proposition to have a high degree of resiliency. Let us say that a language with a probability distribution on it is *really spread out* if and only if every proposition p of the language with positive probability less than one is entailed by a proposition p^* of the language with infinitesimal probability such that the probability of p^* is less than or equal to the probability of not-p. This should not be an extremely difficult requirement for rich languages to meet. Consider the example of the point dart being thrown at the unit interval. For any proper subset S of the unit interval, there will be a proposition of the form that the dart hits point S', where S' is a member of S. It is also the case that the probability that the dart hits S' is less than or equal to the probability that the dart does not land in S; in the extreme case, S will be the unit interval minus one point, and the probability that the dart hits S' will be equal to the probability that it does not land in S. The following theorem places an upper limit on the resiliency of any proposition which has probability less than one:

If a language, together with a probability distribution on it, is really spread out, then the maximum value of the resiliency of a proposition is 0.5, unless its probability is equal to one.

Proof: A preceding theorem showed that if $\Pr(p)$ is not almost equal to one, the resiliency of p had to be equal to zero. Now suppose that $\Pr(p)$ is almost equal to one, but not equal to one. Since the language is spread out, there is a p^* with infinitesimal probability such that p^* implies p. We also know that $\Pr(\bar{p})$ is greater than or equal to $\Pr(p^*)$. But the proposition $p^* \vee \bar{p}$ is consistent with p, and yet the maximum value of $\Pr(p \text{ given } p^* \vee \bar{p})$ is 0.5.

To illustrate this theorem, let p be that the dart does not hit the point $\frac{1}{2}$, $\bar{p}$ be that the dart does hit the point $\frac{1}{2}$, and p^* be that the dart hits point $\frac{1}{3}$. In this case, $\Pr(p^*) = \Pr(\bar{p})$, and $\Pr(p \text{ given } p^* \vee \bar{p}) = 0.5$. These considerations seem to indicate that the only way in which a proposition can be highly resilient is for its probability to be one. This result is unfortunate, because we would like to be able to talk about statistical laws as well as universal generalizations.

After presenting the basic idea of resiliency, Skyrms intends to use it to solve some of the basic problems in philosophy. One problem that he discusses is that of confirmation. It is widely believed that many, if not most, of the laws that we accept are actually false. Thus we find ourselves in the position of wanting to say that a law is well confirmed, and also that its probability of truth is low. A similar problem arose for Carnap, because a consequence of his theory was that every universal generalization had a degree of confirmation of zero. Carnap's solution was to look at instances of the law. Generally speaking, Carnap claimed that what we are interested in is the probability that the next instance we find will accord with the law. For example, suppose our hypothesis is that all ravens are black. What Carnap calls the *qualified instance confirmation* of that law is the degree of confirmation that the next raven we look at will be black. Skyrms believes that this insight of Carnap's is correct, and he uses it in his discussion of the problem of confirmation.

What Skyrms does is give two conditions of adequacy for a measure of confirmation, and then he shows that resiliency is the best measure of confirmation that satisfies these conditions. His first condition of adequacy is:

Criterion I: C(All Fs are Gs) $\leqslant$ Pr(Ga given Fa)(for "new" a)(p. 42).

This criterion tells us that the degree of confirmation of any law must be less than or equal to its qualified instance confirmation. Skyrms considers this requirement very natural; indeed, if it were not true he shows that an instance of a law could disconfirm the law, which is obviously false.

Skyrms' second condition of adequacy is the familiar equivalence condition, which was first discussed by Hempel:

Criterion II: If two laws have logically equivalent associated universal material conditionals, then they are well confirmed to exactly the same degree.

This well known requirement of confirmation also seems to be very natural. The question which then arises is whether any measure of confirmation can satisfy both Criterion I and Criterion II. Skyrms then shows that the only measure of confirmation that could satisfy both of these is one which assigns a zero degree of confirmation to every law. The reason for this is that every law of the form 'All *F*s are *G*s' is equivalent to 'Anything both an *F* and a non-*G* is both an *F* and a non-*F* (p. 50). But the qualified instance confirmation of that is zero: $\Pr(Fa \text{ and } \bar{F}a \text{ given } Fa \text{ and } \bar{G}a) = 0$. Thus no reasonable measure of confirmation can satisfy both Criterion I and Criterion II.

Skyrms avoids this problem by modifying Criterion I in such a way to exclude the unusual laws that raised problems. He does this by first defining a Pickwickian law as one in which the instantiation of the antecedent is incompatible with the instantiation of the consequent. We can see that it was Pickwickian laws which were responsible for our previous troubles. Skyrms then proposes that we replace criterion I with criterion I′:

Criterion I′: If L is non-Pickwickian, $C(L) \leqslant$ Qualified Instance Confirmation (L) (p. 51).

The qualified instance confirmation of a law 'All *F*s are *G*s' is $\Pr(Ga \text{ given } Fa)$. Criterion I′ rules out the types of laws that created problems for Criterion I. Skyrms calls any measure of confirmation which satisfies Criterions I′ and II a measure of convertible confirmation. He then proves that the resiliency of the material conditional $(Fa \supset Ga)$ is the most generous measure of the convertible confirmation of the law 'All *F*s are *G*s'. Thus the notion of resiliency is seen to play a key role in confirmation.

Skyrms replaced Criterion I by Criterion I′ because otherwise no law would have a degree of confirmation greater than zero, which is an undesirable result. Although Skyrms has avoided that result by replacing Criterion I by Criterion I′, we shall now prove that the degree of convertible confirmation of any law is almost zero, i.e., is infinitesimal. Consider the following two laws.

(a) All *F*s are *G*s.

(b) All non-*G*s that are either *F*s or *H*s are non-*F*s.

These laws have logically equivalent material conditionals:

(a′) $(x)[Fx \supset Gx]$

(b′) $(x)[(\bar{G}x \,\&\, (Fx \vee Hx)) \supset \bar{F}x]$.

Thus, by the equivalence condition law (a) is confirmed to the same degree as law (b). Criterion I′ tells us that the degree of confirmation of law (b) is less than its qualified instance confirmation: $C(\text{b}) \leqslant \Pr[\bar{F}a \text{ given } \bar{G}a \,\&\, (Fa \vee Ha)]$. It appears plausible to believe that for all ϵ greater than 0, there is an H such that the probability that a has H is less than ϵ:

$$(\epsilon > 0)(\exists H)[0 < \Pr(Ha) \leqslant \epsilon].$$

This simply tells us that for any real number we pick, there is always some predicate H such that the probability that a has H is less than or equal to that number. We now make use of the following theorem:

$$(\epsilon > 0)(\exists\delta > 0)\,[\Pr(Ha) < \delta \supset \Pr(\bar{F}a \text{ given } \bar{G}a \,\&(Fa \vee Ha)) < \epsilon]\,.$$

This theorem just says that as $\Pr(Ha)$ approaches 0, so does $\Pr[\bar{F}a$ given $\bar{G}a$ & $(Fa \vee Ha)]$. We can use this theorem, and our previous assumption, to get:

$$(\epsilon > 0)\Pr[\bar{F}a \text{ given } \bar{G}a \,\&\, (Fa \vee Ha)] \leqslant \epsilon.$$

This tells us that the qualified instance confirmation of law (b) must be less than or equal to every positive real number. The only possibility is that the qualified instance confirmation must be either 0 or almost equal to zero. By criterion I′ we conclude that the degree of confirmation of law (b) must be either zero or almost equal to zero, and since laws (a) and (b) have logically equivalent material conditionals, we can use Criterion II to conclude that the degree of confirmation of law (a) is also either zero or almost equal to zero. From this we conclude that the degree of convertible confirmation of any universal law is either zero, or differs infinitesimally from zero. Even though Skyrms modified Criterion I in order that the degree of confirmation of a law would not automatically be zero, it appears that the most he gained is that now the law can have an infinitesimal degree of confirmation.

The above problem is the result of Criterion I′ and the equivalence condition. Although I wish to withhold judgement on Criterion I′, I do think that there are some reasons to be skeptical of the equivalence condition. The reason for this is the intimate connection between laws of nature and counterfactuals. The logic of counterfactuals is different from the logic of the material conditional; thus one should not expect two laws which are truth-functionally equivalent to have the same or equivalent subjunctive formulation. For example, consider the two laws:

(c) All *F*s are *G*s
(d) All *F*s are non-*G*s are *G*s

Even though these appear to be totally different laws, they have logically equivalent material conditionals associated with them:

(c') $(x)(Fx \supset Gx)$
(d') $(x)[Fx \,\&\, \bar{G}x) \supset Gx]$

Law (c) may be true, but law (d) seems to be false. It is possible that law (c) supports counterfactual claims, but this does not seem to be true of law (d). It seems plainly false to say that if anything were an F and a non-G, then it would also be a G. This shows that two universal material conditionals being equivalent does not mean that they are associated with the same law. Although the temptation to do so is almost irresistible, it appears that treating laws as material conditionals is an oversimplification. The application of the idea of resiliency to the problems of confirmation is an interesting and promising idea; however a mistake is made when one attempts to treat laws as material conditionals in order to develop this idea.

The purpose of this article has been to discuss critically some of the problems that arise out of Skyrms' discussion of resiliency. The intuitive idea behind the idea of resiliency appears to be very useful, and I suspect that it will play a large role in future philosophy of science. However, Skyrms' present formulation of the principle of resiliency is defective, and needs to be improved.

NOTES

* The author would like to thank John Pollock and Keith Lehrer for numerous insightful discussions which led to this paper, and Clark Glymour for helpful comments on an earlier version. This research was supported in part by Faculty Research funds granted by the University of California, Santa Cruz.

[1] Brian Skyrms: 1980, Causal Necessity, A Pragmatic Investigation of the Necessity of Laws (Yale University Press, New Haven and London). References to this work will be denoted in the text by page number.

Cowell College,
University of California,
Santa Cruz, CA 95064,
U.S.A.