WESLEY C. SALMON*

PROPENSITIES: A DISCUSSION REVIEW

D. H. Mellor, *The Matter of Chance*, Cambridge University Press, 1971, xiii + 190 pp.

Rudolf Carnap and Hans Reichenbach, the original founders and co-editors of *Erkenntnis*, were both intensely interested in the concept(s) of probability, and both made major contributions in this area. At the mid-point of the present century, Carnap's monumental *Logical Foundations of Probability* was published (1950) in the year immediately following the appearance in English of the second revised edition of Reichenbach's *The Theory of Probability* (1949). These two works constitute a bench mark for subsequent work in the philosophy of probability.[1] As even the most casual student of the subject knows, Reichenbach defended an exclusively frequentist theory of probability, while Carnap insisted upon the need for two distinct concepts – degree of confirmation (probability$_1$) and relative frequency (probability$_2$). In the decade following 1950, two new kinds of probability, personal probabilities and propensities, emerged into prominence, and since then they have received considerable attention.

D. H. Mellor's *The Matter of Chance* provides an excellent point of departure for the discussion of these developments, for he attempts to join the two new concepts in a unified theory in ways which are, *with certain important qualifications*, analogous to the manner in which Carnap related his two probability concepts. For Carnap, it will be recalled, degrees of confirmation provide the best estimates of relative frequencies; for Mellor, "To apply the concept of chance to the situation warrants certain partial beliefs on the occurrence of these possible events" (p. 58). Although chance is not identical with propensity, "Chance distributions display dispositions called 'propensities'" (p. 63). In what follows, I shall discuss Mellor's account of the nature of propensities, how they are related to chances, and how chances are related to warranted partial beliefs. For one thing, as we shall see, in Mellor's theory – in contrast to other versions of the propensity theory – propensities are not probabilities, though chances are, as Mellor says on page 1, statistical probabilities. In

Erkenntnis **14** (1979) 183–216. 0165–0106/79/0142–0183 $03.40

the same place, he emphasizes his opinion that statistical probabilities are not to be identified, as Carnap did, with relative frequencies. Nevertheless, in drawing a parallel between Mellor and Carnap, chance and propensity, which are objective and physical, clearly play the role of Carnap's probability$_2$, which shares these characteristics. Another point we must carefully note is that Mellor adopts the unusual strategy of basing his "account of objective probability on a concept of partial belief" (p. 2), instead of going at it the other way around. For this reason, I shall begin by considering the relationship between probability and partial belief.

1. REASONABLE PARTIAL BELIEF

In the history of probability theory, there is a strong tradition linking probabilities with degrees of belief.[2] Within this tradition we find a range of views extending from a purely subjective identification of probability with *actual* degree of belief to a strongly objective requirement that probability be identified with *rational* degree of belief which can somehow be logically justified. As Carnap has pointed out, however, probability theorists sometimes use locutions which make their views sound more subjectivistic than they really intend (1950, §12).

Modern personalism, which might be called "the new subjectivism", derives its chief impetus from L. J. Savage's *Foundations of Statistics* (1954).[3] Recognizing that raw subjective degrees of belief may not even constitute an admissible interpretation of the probability calculus, the personalist imposes conditions of coherence which are tantamount to the requirement that systems of beliefs must satisfy the relations demanded by the probability calculus. Systems of beliefs which do not fulfill these coherence requirements can be shown to be irrational in the sense that, if they are adopted as betting quotients, their holder can be made the victim of a Dutch book – that is, a system of bets which are guaranteed to yield an overall loss no matter what the outcome of the events upon which the wagers are placed. Such coherence requirements are widely regarded as minimal necessary conditions of rationality, but many authors – including Carnap in particular – were unable to accept the official personalist claim that they are also sufficient conditions for rationality of degrees of belief.[4]

In 'The Aim of Inductive Logic' (1962), Carnap explicitly discusses the relationships between actual degrees of belief and rational degrees of belief in terms of credence functions and credibility functions, actual and rational. He explicitly lays down the personalistic coherence conditions as initial rationality requirements. He then goes on to add further axioms, which he spelled out more fully in his 'Replies and Systematic Expositions' (§25–26) in Schilpp (1963). His full list of axioms for inductive logic contains a motley assortment of requirements on confirmation functions or rational credibility functions, whose justifications – if indeed they have any – seem odd, various, and unclear. In the end, Carnap makes his final appeal to "inductive intuition". Carnap is, it seems to me, clearly correct in seeking rationality requirements which go beyond mere coherence, but his attempt to justify the further conditions is patently inadequate. Something more is needed, and Mellor's attempt to provide a way of warranting partial beliefs may be a step in the right direction.

Partial beliefs are, according to Mellor, warranted by chances, and chances are physical probabilities. They are, in fact, precisely those features of the world which provide the warrant for partial beliefs. "What makes [a chance statement] true is a correspondence between the partial belief it expresses . . . and some objective feature of the world" (p. 27). Unlike Carnap, Mellor makes no attempt to provide an "inductive logic" or "theory of confirmation" to enable us to see how chances are to be ascertained, or how to confirm the statistical laws which express chances. Roughly speaking, Mellor tells us that, if we know what the chances are, and if we so fashion our partial beliefs that they agree with the chances, then our partial beliefs are warranted. More precisely, if we assign chances to events whose numerical values are identical with the warranted partial beliefs in these events, then we have assigned the chances correctly. Either way, we find a strong disanalogy between Mellor's union of the two types of probability and Carnap's. This disanalogy has deep philosophical import; we must examine it with some care.

In *Logical Foundations of Probability*, the singular predictive inference – i.e., the inference from some observed sample of a population to an unobserved member of that population – played several important roles. First, the degree of confirmation of the statement that individual i has property F provides the *best estimate* of the relative frequency with which property F occurs in the population at large. Second, the same degree of confirmation furnishes the *fair betting quotient* upon which to base a wager on whether i has property F. Third, the singular predictive inference gives us *qualified-instance confirmation*, which is Carnap's surrogate for the confirmation of general hypotheses. The fact that Carnap's system yields degree of confirmation zero (in L_∞) for all universal generalizations on any finite amount of evidence forces him, in effect, to substitute the concept of the fair betting quotient on the next instance for the notion of confirmation of a law-statement. This feature of Carnap's theory of confirmation which has been circumvented by Hintikka (1965), is not acceptable. If the degree of confirmation of the statement, "All beings are imperfect", is zero, then the statement, "There exists a perfect being", has degree of confirmation 1. This result might provide some solace to the followers of Anselm and Descartes, but it is unacceptable in a theory of scientific confirmation. Carnap's system fails altogether to deal with the confirmation of scientific hypotheses – a task Mellor explicitly declines to treat (p. xi).

Leaving that issue aside, Carnap's approach to inductive logic runs into serious trouble on account of his identification of the best estimate of the relative frequency with the fair betting quotient. A rational person, faced with a possible wager, should ask two distinct questions: (1) With what relative frequency will events of the type on which I intend to bet occur within some appropriate class? (2) To what extent can I rely upon my answer to question (1)? Since, in most cases of interest, we cannot know with certainty the relative frequency, we must use, as Carnap would say, the best estimate. According to Carnap, degree of confirmation statements are analytic (if true), and they provide what is, by definition, the best estimate. Hence, although we cannot be certain of the value of the relative frequency, we can be certain of the best estimate of the relative frequency. And this, Carnap claimed, is precisely the fair betting quotient.

These considerations lead to two severe difficulties. First, there is the so-called "paradox of perfect evidence". Suppose we are about to draw a ball from an urn, and we are contemplating a bet on its color. It is easy to arrange for the *a priori* probability (on no observed instances) of red on the forthcoming draw to be 1/2. It is also easy to design a very large sample from this urn so that the *a posteriori* probability of red on the next draw, given an observed sample of just this sort, is also 1/2. Nevertheless, it is evident that these two "best estimates", which are identical in numerical value, are *not* equally good for purposes of making bets.

Although it might be reasonable to use either of them as a betting quotient for a small wager, it would clearly be foolhardy to use the *a priori* probability as a basis for wagering a large portion of one's total fortune. Nevertheless, it might conceivably be rational in some circumstances to use the *a posteriori* probability founded upon a massive body of evidence as the basis for a large wager. The point is, of course, that one cannot simply take numerical values of probabilities as fair betting quotients in order to determine betting odds for wagers of stakes of all sizes. I am not referring to the diminishing marginal utility of money, which, as Carnap himself emphasized, must be taken into account. Even if all wagers are measured in units of utilities rather than monetary value, it is still unreasonable to allow bets to be determined solely by Carnapian fair betting quotients.

The second difficulty results from the fact that 1 is never a reasonable choice for a betting quotient. Since a betting quotient is defined as the individual's stake u_1 divided by the sum of that stake and the opponent's stake $u_1 + u_2$, a betting quotient of 1 would license a wager of an arbitrarily large sum, say 10^9 DM, that the event in question will occur against a stake of zero that it will fail to occur. This problem affects Mellor's theory:

> . . . a deterministic law ascribes a chance of 1 to an occurrence of one event (of kind *B*, say) given the occurrence of another (of kind *A*). What does this amount to on the present theory? We have that, assuming the law, upon each occurrence of an *A*-event the reasonable partial belief to adopt on the occurrence of a *B*-event is of degree 1. This entails that it is in these circumstances unreasonable to put any money on a *B*-event not happening, whatever odds are offered. . . . If I know that every *A*-event, past, present, and future, is accompanied by a *B*-event, I know that this one is. It would be unreasonable of me in this assumed state of knowledge to put any money on the *B*-event not happening, since I know in advance that I would lose it (p. 159).

In a certain sense, Mellor's argument seems eminently sensible, but upon slight reformulation, a difficult problem emerges. If, in some acceptable sense of the word "know", I know that all *A* are *B*, would it be rational to accept a bet of one pfennig that the next *A* will be a non-*B* if I am required to put up 10^9 DM on the opposite side? Or more poignantly, should I be willing to put up 10^9 DM against a grain of rice, or even against nothing at all? I doubt that "knowledge" of this kind ever exists.

One might suppose that bets of this sort are not really at issue, but if we take Mellor at his word, they are. The betting quotient is subject to certain constraints: "(i) the gambler has to bet, (ii) he chooses the betting quotient, and then (iii) his opponent chooses the stake size and the direction of the bet" (p. 161). These constraints "are needed to ensure that the choice of betting quotient measures the gambler's partial belief and nothing else" (p. 160–161).

There are, it seems to me, many occasions on which a reasonable estimate of a relative frequency is one – e.g., the frequency with which objects made of metallic copper are good conductors of electricity. I am not sure whether this estimated value of the relative frequency is a good value for a degree of belief, for there are two distinct issues. The first question concerns the most reasonable estimate of the relative frequency. The second question concerns the degree of certainty that the estimate matches the true value. Both considerations have to enter the betting situation. It seems clear to me that the best estimate of the relative frequency cannot be taken as the correct betting quotient in all cases. Let me distinguish two concepts:[5] (1) the *fair* betting quotient is the betting quotient which is numerically equal to the relative frequency of the outcome in the class in question; in a sense to be discussed more fully below, it gives neither bettor an advantage in a gamble. (2) The *rational* betting quotient is the most reasonable determination of odds in terms of the knowledge available to the gambler if he is forced to bet on terms outlined by Mellor. The point is this. Since we do not know for certain the value of the relative frequency, and therefore, the value of the *fair* betting quotient, we must on various occasions try to arrive at reasonable estimates of them. There are many occasions on which a good estimate of the relative frequency is distinctly not a good estimate of the fair betting quotient, if by "a good estimate of the fair betting quotient" we mean a reasonable value to use in making bets – i.e., a *rational* betting quotient. The moral is very simple: what constitutes a good estimate of a given quantity for some purposes will not always be a good estimate of the same quantity for other purposes. Carnap tried to work out a system of inductive probability in which the two could be identified, but as we have seen, the effort was unsuccessful. I am inclined to think that it is impossible in principle to achieve this goal.

Mellor claims that our partial beliefs are *warranted* if they agree in value with the chances, and I think this point, which was persuasively argued by F. P. Ramsey (1931), is quite correct. Mellor, however, attempts to support it with an argument I find altogether unconvincing. This argument deserves to be examined with some care. "Suppose the frequency of B-events in the sequence of N trials is f. The net profit per unit stake is the difference between the frequency and the agreed CBQ, $|f - r|$" (p. 161). The CBQ is the coherent betting quotient $r = u_1/(u_1 + u_2)$. The net outcome of such a series of bets is that the subject wins fN of the bets, gaining u_2 in each case, and he loses $(1 - f)N$ of the bets, losing u_1 in each case. Thus, the total gain G (which may assume a negative value) is given by

$$fNu_2 - (1 - f)Nu_1 = fN(u_1 + u_2) - Nu_1 = fN(u_1 + u_2) - N(u_1 + u_2)\frac{u_1}{u_1 + u_2} = N(u_1 + u_2)\left(f - \frac{u_1}{u_1 + u_2}\right) = N(u_1 + u_2)(f - r).$$

Thus, when Mellor refers to the "net profit per unit stake", he is referring to the total stake $N(u_1 + u_2)$ wagered by both players in the aggregate of all of the plays. If, for example, the odds are 2:1, one player betting 10 DM against the other's 20 DM on 100 plays, the total stake is 3000 DM. Since the total stake gets large as N increases, the assurance that $f - r$ gets small under certain conditions does *not* imply that the total profit or loss G will be small in a large number of trials.

Mellor appeals to the strong law of large numbers, which tells us "that there is an arbitrarily high probability $1 - \varepsilon$ that after some sufficiently large number of bets, N, the frequency f will differ from p [the probability] by less than some arbitrarily small amount δ" (p. 161). This result is needed "to show how that law can provide knowledge that a gambler will break even" (p. 160). "The arbitrarily small constant δ plainly explicates the concept of breaking even. The gambler breaks even if and only if the net profit on the sequence is less than δ" (p. 162). These statements can be considered technically correct only if certain key terms, such as "net profit" and "breaking even", are construed in rather unusual ways. Such statements are, consequently, open to easy misinterpretation. It is essential to make clear, in particular, that when Mellor speaks of "the net profit" he is referring to the *average* net profit *per play*. Some people might be quite unhappy with his use of terms. If one plays a sufficiently large number of times – and note that the law of *large* numbers refers to some sufficiently *large* number of bets N – then a small loss (negative profit) per trial on the average is completely compatible with a very large overall loss. Suppose a man goes out one evening and plays the standard game of heads and tails at even odds with a stake of 1 DM per trial. He plays 100,000 times, winning 49,500 times and losing 50,500 times. His betting quotient

$r = 0.5$, and the frequency $f = 0.495$; his δ can thus be set at 0.005. His average loss per trial is 1 pf. When he comes home and his wife asks how he made out, he reports that he "broke even" (just about) and that his "net loss" was 1 pf. She breathes a sigh of relief. "At least you didn't lose your shirt as you usually do". I shall refrain from quoting her remarks upon learning that he came home from the game 1000 DM poorer than when he left.

Since the foregoing argument plays a crucial role in Mellor's theory of chance, let us take a closer look at what happens to the fortune of a player of the standard game. We assume that the odds are 1:1, and that the probability of heads is, in fact, 1/2. On each trial, the player either gains or loses 1 DM, so his total fortune changes by either $+1$ or -1, and on each trial the probability is 1/2 for each alternative. Thus, his fortune, which starts from the initial value F_0 and assumes the value F_n after the nth trial, executes a one-dimensional symmetric random walk.[6] In the long run, it assumes every finite value an infinite number of times, so his gains and losses become arbitrarily large, but it returns to the initial value F_0 infinitely many times, so every loss or gain is temporary. There is, of course, no guarantee that F_n will return to its original value by the Nth trial. If $r = p$, the law of large numbers tells us that the difference between f and r will be small after a sufficiently large number of trials N, but it does not in any way imply that the difference between the initial fortune F_0 and the fortune after the Nth trial F_N – that is, the total net gain or loss G – will be small. For sufficiently large N, an arbitarily small difference δ between f and r is compatible with an arbitrarily large (positive or negative) gain G. A small value of δ therefore seems to be a funny sort of "breaking even". In this sense, breaking even simply means not losing your shirt too fast.

Mellor's analysis of the rationality of a betting quotient which is equal to the probability of the occurrence upon which the wager is being made thus appears seriously defective. This is not to say that his conclusion is false, for certainly there is a very special relationship between values of objective probabilities and rational betting behavior. The failure of Mellor's argument reflects the very profound philosophical problems associated with the application of knowledge of probabilities to finite sets of events. The inadequacy of this argument also seriously undercuts Mellor's strategy of basing "an account of objective probability on a concept of partial belief" (p. 2).

Since the type of argument Mellor uses to show the unique rationality of a betting quotient which matches the probability of the event in question has a strong intuitive appeal, it might be supposed that the argument could somehow be reconstructed to avoid the foregoing objections. I should like therefore to take a closer look at the strong law of large numbers, to which Mellor makes fundamental appeal, and at the use he makes of it. Two points require explicit comment. First, the law to which Mellor appeals has an antecedent condition which he fails to mention, namely, that the events involved are *independent* trials. In any concrete situation, it is an empirical matter to determine whether this condition is satisfied or not. We know of many sorts of examples, such as coin flipping and spontaneous radioactive decay, in which the condition holds; we know of many others, such as induced fission of U^{235} and learning with positive reinforcement, in which it does not. There is a serious risk, it seems to me, in taking a law which is not applicable to all probabilistic situations in general as the foundation of a philosophic theory of probability.

Second, since the very meaning of "probability" is at issue, we must look closely at the way in which this term occurs in the law of large numbers. Mellor says:

> The gambler, we suppose, holds there to be a chance p of a B-event on each occasion of the bet. If he is right then the strong law of large numbers shows of p alone that there is an arbitrarily high probability $1 - \varepsilon$ that after some sufficiently large number of bets, N, the frequency f will differ from p by less than some arbitrarily small amount δ. This is on the present account to say that some partial belief of strength p, and no other, is such that an arbitrarily strong partial belief $1 - \varepsilon$ is reasonable on the bet that $|f - p| < \delta$ (pp. 161–162).

The concept of probability occurs in this passage twice, and at two distinct levels. The chance p is, on Mellor's terminology, the statistical probability of B in this context. The law asserts a relationship between this probability p and the frequency f of B in a sequence of N trials. The relationship asserted is probabilistic; it involves the probability $1 - \varepsilon$. The question is, what meaning are we to assign to this second level probability? Mellor refers to it as "an arbitrarily strong partial belief [which] is reasonable", but what can that mean? As he explicitly acknowledges, he cannot interpret this probability as chance:

> But the law refers in turn to a CBQ, $1 - \varepsilon$, being reasonable on the many-case bet. This reference must be accounted for without further appeal to chance, if our explication of the single case is not to be viciously circular (p. 162).

He certainly does not want to construe it to mean that this CBQ (coherent betting quotient) will yield a high percentage of successful bets, for that would involve him in a frequency conception of probability, which he is at great pains to reject.

Mellor must somehow show that it is reasonable to rely on high probabilities for betting purposes without tacitly supposing that such bets, by definition, will usually be won.[7] Since I find it difficult to understand Mellor's attempt to deal with this issue, let me quote his answer in full:

> I rely here on an asymptotic property of very high reasonable CBQs. The property is roughly that the better a very good bet is, the more circumstances there are in which the gambler may act as if he knew it was won. In chapter I I argued that increasing partial belief must merge into full belief. In particular (pp. 6, 15), as one's partial belief in a proposition increases, it must in the end amount to belief in the proposition. So a reasonable degree of partial belief which tends to 1 implies eventually that the corresponding full belief is reasonable (p. 162).

My difficulty can be posed in the following way. Suppose the circumstances are such that for a given δ and N, $\varepsilon = 1/2$, and so $1 - \varepsilon = 1/2$. For the same value of δ but a larger N, we have $\varepsilon' = 1/4$. For still larger N, we have a series of values ε'', ε''', ..., converging to 0. Now it seems to me that Mellor's argument tells us that *if* $1 - \varepsilon$, $1 - \varepsilon'$, ... are *reasonable* degrees of partial belief, then with ever decreasing values of ε we are approaching full *reasonable* belief. But I cannot see that he has offered any non-question-begging reason for saying that any of these second level probabilities mentioned in the law of large numbers represents a reasonable degree of belief. And in the absence of such assurances, the fact that we have a sequence of probabilities converging to 1 does not seem to have much bearing upon what it is reasonable to believe. Calling $1 - \varepsilon$ a CBQ goes no farther than asserting that to regard it as a probability will not lead to incoherence. That falls far short of demonstrating that it is a reasonable degree of partial belief.

2. PROPENSITIES

Let us now turn our attention more specifically to physical probabilities. At mid-century the dominant concept of physical probability was embodied in the limiting frequency interpretation, but before the end of the following decade, Karl Popper, who had previously been a frequentist, began advocating the "propensity interpretation" in a series of influential articles (1957, 1960, 1967). This view has subsequently attracted a good deal of attention and a number of strong adherents – Mellor being one of the most prominent (though it is more accurate to say that he adopts a propensity *theory* of probability rather than a propensity *interpretation*, since propensities are not probabilities on his view). A comprehensive survey of the major proponents and various versions of the propensity theory has been furnished by Henry Kyburg (1974). Amongst philosophers, it appears that the propensity theory is gaining in popularity, while the frequency theory is losing ground.

Popper's initial presentation of the propensity interpretation was in a brief paper included in a symposium on quantum mechanics. Noting the brevity of this article, Popper promised, "A full treatment of the propensity interpretation and its repercussions on quantum theory will be found in the *Postscript: After Twenty Years* to my *Logic of Scientific Discovery*, 1957" (1957, p. 65, fn.). This book was never published. In the 1957 paper, Popper made some extravagant claims about his new interpretation:

> The main thing about the propensity interpretation is that *it takes the mystery out of quantum theory, while leaving probability and indeterminism in it.* It does so by pointing out that all the apparent mysteries would also involve thrown dice, or tossed pennies – *exactly* as they do electrons. In other words, it shows that the quantum theory is a probability theory just as any theory of any other game of chance, such as the bagatelle board (pin board) (1957, p. 68, Popper's italics).

To take the mystery out of quantum theory would be a neat trick, indeed, but there are few, if any, who think the propensity interpretation does so. Without retracting this claim, Popper does shed doubt upon it in a later article. Speaking of the pin board, he remarks parenthetically, "This has its similarity with the two slit experiment, even though we have here no superposition of amplitudes..." (1967, p. 33). A little later, still referring to the pin board, he says, "There will be no interference of amplitudes: if we have two slits Δq_1 and Δq_2, the two probabilities themselves (rather than their amplitudes) are to be added and normalized: we cannot imitate the two slit experiment" (1967, p. 34). Even more explicitly, he adds, "The peculiarity of quantum mechanics is the principle of the superposition of wave amplitudes – a kind of *probabilistic dependence* (called by Landé '*interdependence*') that has apparently no parallel in classical probability theory" (1967, p. 40, Popper's italics). Most theorists would agree, I think, it is the very superposition of amplitudes which constitutes one of the major mysteries of quantum theory. If there is, in fact, "no parallel in classical probability theory", then there is *no way* to show "that the quantum theory is a probability theory just as any theory of any other game of chance". Whatever the other virtues of the propensity interpretation, it does not take the mystery out of quantum theory. I shall not discuss this issue further; instead, I shall concentrate upon the propensity interpretation as such, considering it on its own merits.

In order to introduce the propensity interpretation, Popper invites us to consider a simple situation, namely, a sequence of throws of a loaded die for which the probability of side six showing is 1/4. If this die were tossed repeatedly in the standard manner the relative frequency of six would converge to the limit 1/4. He then poses the following "awkward question":

> What if the sequence consists of throws of a *loaded* die, with one or two throws of a *regular* die occurring in between the others? Clearly, we shall say about the throws with the regular die that their probability is different from 1/4, in spite of the fact that these throws are members of a sequence of throws with the frequency 1/4.
>
> This simple objection is of fundamental importance. It can be answered in various ways. I shall mention only two of these answers, one leading to a *subjective interpretation*, the other to the *propensity interpretation* (1957, p. 66, Popper's italics).

Although Popper acknowledges that the frequentist can go some way in answering this objection, by pointing out that the tosses of the standard die could be referred to a different (possibly hypothetical) sequence, namely, tosses of that standard die, he does not believe that the statistical interpretation (frequency interpretation) can provide an adequate answer.

As Popper points out explicitly, he is concerned with the probability of a single event. Although the frequentist Richard von Mises denied that probabilities could meaningfully be assigned to single events, other frequentists – most especially John Venn (1866) and Hans Reichenbach (1949) – offered answers to this problem. Popper seems to have misunderstood their answer. As he characterizes it, "what it *means* to say 'The probability of throwing 6 *with the next throw* of this loaded die is 1/4' ... can only mean one thing: 'The next throw is a *member of a sequence* of throws, and the relative frequency within this sequence is 1/4'" (1957, p. 66, Popper's italics). As Venn had pointed out almost a century earlier, the given single event is a member of many sequences, and the problem is precisely that of picking the correct sequence to which to refer it. The solution, as I prefer to put it, is that the single event should be referred to the broadest homogeneous reference class, and its probability or weight should be taken as the limiting frequency in *that* particular reference class. One does not pick just any sequence which happens to be handy. Clearly, it will not do to refer the one or two tosses with the standard die to the sequence which consists otherwise of tosses of the biased die.

It is notable that Reichenbach had addressed precisely the problem Popper poses:

It is a consequence of the [frequency definition of probability] that the degree of probability is considered as a property of a sequence in its entirety. In many applications, however, we deal with sequences for which we want to express the fact that a definite probability exists for each individual element, that is, the probability is constant from element to element. In order that such a statement will not contradict the general logical structure of the probability concept, we must investigate how the statement about an individual element of a probability sequence can be translated into a statement about a whole sequence – that is, in the language of the frequency interpretation, how it can be expressed as a statistical statement (1949, pp. 167–168).

Popper argues that it is not enough to look at frequencies in sequences; "we consider as decisive the *conditions under which the sequence is produced*" (1957, p. 67, Popper's italics). Reichenbach had said:

When we produce a probability sequence by throwing a die, we demand that each throw be played with the same probability, that is, there should not be occasional exceptions where a loaded die is used or the die is insufficiently shaken. The requirement signifies a well-defined condition for the physical production of the sequence... For example, if we do not make the fourth throw properly, but produce it by deliberately placing face 6 up, the incorrectness will not show in the limit of the frequency, because a single throw does not alter the limit properties of the whole sequence (1949, p. 168).

In order to deal with this situation, Reichenbach introduces what he calls a "probability lattice" – an infinite sequence of infinite sequences. In this lattice we can examine the limiting frequencies in vertical columns as well as horizontal rows.

Imagine, for instance, that all the horizontal sequences contain, in the fourth throw, one result obtained by placing the face 6 up; this would show in the lattice by the fact that the fourth vertical sequence possesses for the "6" the frequency 1 as its limit....

Reversing the inference, we can express the assumption that the probabilities within the horizontal sequences are the same from element to element by postulating that each vertical sequence should likewise possess the limit p for its frequency (1949, p. 169).

Reichenbach does not, of course, postulate the physical existence of infinite sequences of infinite sequences. All of the leading frequentists from Venn on down have recognized that the infinite sequence is a mathematical idealization; it represents, in some sense, what would have happened if an infinite number of trials had occurred. This may or may not be objectionable, but in any case, it is a standard part of the frequency theory.

It has often been noted that Popper seems to offer two distinct versions of the propensity interpretation – a "virtual sequence propensity interpretation" and a "single case propensity interpretation".[8] The former version is found in such passages as the following:

Every experimental arrangement is *liable to produce*, if we repeat the experiment very often, a sequence with frequencies which depend upon this particular experimental arrangement. These virtual frequencies may be called probabilities. But since the probabilities turn out to depend upon the experimental arrangement, they may be looked upon as *properties of this arrangement. They characterize the disposition, or the propensity*, of the experimental arrangement to give rise to certain characteristic frequencies *when the experiment is often repeated* (1957, p. 67, Popper's italics).

The frequentist insists that a probability represents a relation between two classes – a reference class and an attribute class – or two sequences. In the die casting case, the reference class consists of the series of throws of the die, while the attribute class consists in landings with the side 6 uppermost. Indeed, in Reichenbach's full formulation, we begin with two ordered sequences $\{x_i\}$ and $\{y_i\}$, the members being coordinated via the subscripts. For the sake of simplicity, we may assume that every x_i belongs to the reference class A; consequently, each x_i would be a toss of the die. Each y_i would be an event of the die coming to rest on the surface on which it is thrown; those cases in which the side 6 is on top belong to the attribute class B. Clearly the definition of the attribute class A specifies an experimental arrangement which permits repetition of the experiment. As long as the reference class is clearly defined, we can say that each of the trials is made under the same experimental arrangement. Up to this point, I can find no difference in substance – only terminological differences – between the virtual sequence propensity interpretation and the standard frequency interpretation. When Popper says that "the propensity interpretation of probability. . .differs from the purely statistical or frequency interpretation only in this – that it considers the probability as a characteristic property of the experimental arrangement rather than as a property of a sequence" (1957, pp. 67–68), it seems fair to say that the difference is, at most, one of emphasis. Most current propensity theorists would, I believe, reject this explication in terms of virtual frequencies and virtual sequences as a dispensable vestige of Popper's former frequentism. The full-blown single case propensity interpretation enjoys much greater favor these days, and Mellor's book reflects this point of view.

The attempt to provide an interpretation of probability which applies directly to single events is clearly (despite possible lapses) one of Popper's basic motivations in developing the propensity interpretation:

> The main point of this change is that we now take as fundamental the probability of *the result of a single experiment*, with respect to its *conditions*, rather than the frequency of results in a sequence of experiments. Admittedly, if we wish to *test* a probability statement, we have to test an experimental sequence. But now the probability statement is not a statement *about* this sequence: it is a statement *about* certain properties of the experimental conditions, of the experimental set-up (1957, p. 68, Popper's italics).

One of the major problems to which this approach gives rise is, it seems to me, just the same as it is for the frequency interpretation – namely, the problem of uniqueness. This is the problem from which Popper's discussion takes its departure – how can we assign a probability other than 1/4 to the result 6 *when the unbiased die is thrown*? Popper's chief criticism of the frequentist position is that it seems to lead to two distinct values, depending upon which sequence you choose as the reference class. Now, Popper tells us, the single case probability statement is "about certain properties of the experimental conditions". But which among *all* the properties of an experimental set-up are the "certain properties" to which the probability is relative? Are we talking about the toss of any old die, biased or unbiased? or about any method of tossing? or tossing by means of a dice cup? or by a left-handed person? or from a certain height above the table? or with a closely specified angular momentum? It seems to me that the only suitable answer available to Popper is to say that we must consider the *relevant* properties of the experimental set-up. Notice the similarity between this answer of the propensity theorist and the answer given by the frequentist. The frequentist (this frequentist, at any rate) says that the probability assigned to the single case must be based upon the relative frequency in the broadest homogeneous reference class – that is, the class defined in terms of all and only those factors which are statistically relevant to the occurrence of the attribute in question. For a frequentist, a feature of the experimental set-up is statistically relevant if and only if it makes a difference to the long run relative frequency. For a propensity theorist such as Popper, a factor would be statistically relevant if and only if it makes a difference to the strength of the propensity to produce the outcome in question (see Fetzer, 1974, p. 392).

3. CHANCE

Mellor's method of coping with this issue is rather different from Popper's. At the outset of his discussion, Mellor states explicitly that he is concerned to provide a philosophical account of statistical probability – which he calls 'chance' – and which is not to be identified with relative frequency. This concept, he claims, must endow chance with four characteristics: it is (i) objective, (ii) empirical, (iii) non-relational, and (iv) applicable to the single case. "The chance of a radium atom decaying in the next ten minutes is as objective and empirical a matter as its mass, as little relative to evidence, and as much an attribute of one as of many statistical trials" (p. ix). Let us accept for the sake of discussion the first two characteristics, agreeing that the sort of physical probability we seek is objective and empirical. The remaining two characteristics pose the major problems. There is a clear sense in which we can agree with Mellor that the physical probabilities with which we are concerned should not be relative *to evidence*, as Carnap's logical probabilities are, but it is not so clear that they cannot be relational in any other sense. The frequentist, who is just as concerned as Mellor to have probabilities which are objective and physical, still maintains that the probability of a given outcome is relative to a set of physical circumstances, and on Popper's version of the propensity interpretation, propensities also appear to be relative to physical conditions. But Mellor will have none of this (see p. 60). Instead of saying, as Popper did, that probability statements are about *certain* features of the experimental situation, Mellor seems committed to say that they must be about *all* of the characteristics of an object – about the object in its full individual particularity. If a coin, tossed repeatedly by a certain mechanism, lands heads up about half of the time, that does not mean that the chance of heads is 1/2. Suppose for the moment that determinism is true. If we look at the particular toss in great detail, ascertaining precisely such quantities as the angular momentum imparted to the coin – and many others as well – then we would find that the coin *must* land heads up, and the chance of heads would be not 1/2 but 1. Even if determinism is not strictly applicable to such situations, the probability would be close to 1 if not precisely equal to 1. We are certainly prohibited, on Mellor's theory, from saying that the probability of heads is 1/2 relative to tossing in a certain incompletely specified manner; when *all* things are taken into account the chance is one. It follows, then, that if determinism is true, there are no chances (other than 0 or 1). "Chance distributions display dispositions called 'propensities'" (p. 63). "If propensities are ever displayed, determinism is false" (p. 151).

This consequence of Mellor's theory is, to my mind, unacceptable, and quite unnecessary in an adequate theory of chance. Whether determinism is true or false, we must be prepared to assert that there is a chance of 1/2 that a particular coin tossed in a certain way will come up heads, a chance of 1/2 that a certain molecule will be located in the left half of a container of gas at some particular time, a chance of 1/2 that a person who smokes cigarettes at a certain rate for a certain length of time will contract lung cancer, a chance of 1/2 that a given U^{238} molecule will spontaneously disintegrate in 4.5×10^9 years, and a chance of 1/2 that a person in a particular mental state will make a suicide attempt within a week. To suppose that the truth of such statements depends upon the truth of indeterminism flies in the face of virtually universal common and scientific usage. If chance is identified with statistical probability, and chance does not exist if determinism is false, then the truth of all statistical laws depends upon the truth of indeterminism. To make sense of statistical sciences – and which sciences are not statistical in some respects? – we need to allow that chances, though admittedly not relative to bodies of evidence, are relative to specifiable physical features of experimental set-ups. To claim, as Mellor does, that statistical laws would, under deterministic conditions, express useful fictions, does not, I believe, do adequate justice to statistical laws.

4. PROPENSITIES AND CHANCES

I remarked above that for Mellor propensities are not probabilities.[9] It is time to elaborate that point. Mellor maintains that there are objects such as dice, radioactive atoms, and human beings which have certain dispositions known as propensities. When an object has a dispositional property, such as fragility or hostility, then it displays the disposition when suitable circumstances arise. A fragile glass breaks when dropped on a hard floor; a hostile person manifests angry behavior when provoked. For Popper, and many other propensity theorists, a coin has a propensity of a certain degree or strength *to land heads up* when tossed in the standard way. The propensity to show heads, on this account, manifests itself about half of the times the coin is tossed. Such dispositions may differ in this one important way from dispositions of the more usual sort, but that does not make them impossible or unintelligible. Mellor does not agree with this sort of analysis; he maintains that a propensity is a disposition to *display a chance distribution*, e.g., the coin, when tossed, *always* displayed a chance of 1/2 for heads and 1/2 for tails. Since the propensity is constantly present, and displays itself on every suitable occasion, it does not admit of degrees or strengths which could assume values satisfying the axioms of the probability calculus. Propensities are not chances, and it is the values of the chances, not the values of propensities, which constitute an interpretation of the probability calculus.

As Mellor admits (p. 65), it would not be altogether contrary to accepted usage to allow, for example, that a person of generous disposition might on rare occasions behave meanly, but to avoid ambiguity, he introduces a sharp terminological distinction. "Hereafter I intend 'disposition' to imply invariably in this sense, and use 'tendency' for what can admit unexplained exceptions to the regularities of behavior it purports to explain" (p. 65). According to this terminology, Popper (along with many other propensity theorists) regards propensities as tendencies which admit of degrees or strengths, while Mellor sticks to the claim that propensities are physical dispositions. He supports his position on the basis of the following argument:

> We face the following dilemma. Either a propensity is not a disposition or results and outcomes of chance trials do not display it. Let us test the first horn of the dilemma. The result of a chance trial could be taken to display what I have called a 'tendency'...
>
> What is wrong with...this, of course, is that the concept of [tendency] needs analysis at least as much as chance does. And, moreover, no analysis seems plausible that does not involve essential reference to chance or...reduce essentially to relative frequency...The former would be viciously circular and the latter simply amounts to a frequency account of chance that we have already rejected.
>
> I settle therefore on the other horn of the dilemma, taking propensity to be a disposition and denying that the result of a chance trial is its display (pp. 69–70).[10]

Thus, to reiterate, the propensity is a disposition to display a *chance* *distribution*, such as {chance of heads = 1/2; chance of tails = 1/2}, rather than a tendency to display a *result*, such as heads.

One possible objection to Mellor's position is that we can see that a coin comes up heads on some occasions, and so we can see when the propensity manifests itself and when it does not on the usual conception. On Mellor's analysis, it might be asked, how can we see that the propensity is manifesting a chance distribution on a single toss? What we can see is that the result is a head or a tail. In what sense can a chance distribution be *displayed*? To such objections Mellor appropriately replies that not all dispositions are dispositions to display directly observable results. A person may have a disposition to become unconsciously angry whenever he encounters someone who resembles his father, but this anger may not be *directly* observable to the subject or anyone else. An atom in a crystal may have a disposition to vibrate with greater displacement from its position in the lattice when the crystal is heated, but such agitation is not *directly* observable. This response does not, however, release Mellor from the obligation of providing physical meaning for the concept of displaying a chance distribution as it applies to a single trial. As we have seen, Mellor's way of doing so involves an explication of chance distributions in terms of warranted partial beliefs, and we have seen that his key arguments are, at best, dubious.

Let us, therefore, consider the grounds on which Mellor rejects the first horn of his dilemma – in particular, his arguments against the frequency theory. This issue is crucial, for the propensity theory in all of its various forms has been widely touted as a significant improvement over the long-accepted frequency conception. If it is, indeed, such an improvement, we owe it to ourselves to see precisely in what respects it enjoys its superiority.

At the outset, Mellor expresses his view that the "[f]requency theory makes no sense of the single case..." (p. xi). Later on, he adds, "Single case application is so central to the use of probability statements...that no account of their meaning can be acceptable which rules it out" (p. 54). This objection cuts sharply against any frequentist of the von Mises variety who asserts, "We can say nothing about the probability of death of an individual...It is utter nonsense to say, for instance, that Mr. X, now aged forty, has the probability 0.011 of dying in the course of the next year" (quoted by Mellor, p. 53). But other frequentists, such as Venn, Reichenbach, and I, have tried seriously to come to grips with the problem of the single case.[11] Mellor gives these efforts short shrift; all are dispatched in a single paragraph (p. 54):

This view [the frequency theory] makes it quite inexplicable how statistical data can be given individual application. Reichenbach (1949, §71) rightly insists that "it is the *predictional value* that makes probability statements indispensable" but is reduced to the following account of how this indispensable rôle can be fulfilled (§72):

> An individual thing or event may be incorporated in many reference classes, from which different probabilities will result. . . We then proceed by considering *the narrowest reference class for which reliable statistics can be compiled*. . . . We do not affirm that this method is perfectly unambiguous. . . we are dealing here with a method of technical statistics.

Of course we are dealing with nothing of the sort: we are dealing with a crucial test of the adequacy of any proposed definition of chance. The same goes for Salmon's (1967, p. 91) alternative appeal to the "*broadest homogeneous reference class*" in providing for the application of statistical data to the single case. It will not do to relegate this "extremely practical affair" (Salmon, 1967, p. 92) of obtaining "weights that can be used in practical decisions" (p. 95) to "rules [that] are part of methodology, not of probability theory" (pp. 93–94). The process of picking particular reference classes is no doubt extremely practical; but if the process is what makes probability statements indispensable probability theory should at least make sense of it. The issue cannot be disposed of just by calling single case chances 'weights' and the problems of determining their values 'methodological'. Single case application is so central to the use of probability statements (cf. Ayers, 1968, p. 23) that no account of their meaning can be acceptable which rules it out.

I have quoted the paragraph in its entirety because of its central role in Mellor's rejection of the frequency theory.

Neither Reichenbach nor I would dissent in the least from the sentiment concerning the importance of applicability of probabilities to single cases expressed in the last sentence. As a matter of fact, in (1967, chap. IV, §2), I set forth three general criteria of adequacy for interpretations of probability, namely, admissibility (satisfaction of the axioms of the probability calculus), ascertainability (possibility in principle of establishing values of probabilities), and applicability (utilizability of probabilities "as a guide of life"). In my discussion of the frequency interpretation, I specifically cite the problem of the single case as the most serious difficulty encountered by the frequency interpretation in connection with the criterion of applicability (1967, p. 90). And I attempt to provide an answer which is appropriate to the frequency interpretation. While it is true that I do not offer a definition of probability which makes that term apply semantically to single cases, I do go to some lengths to show how knowledge of probability values has practical and theoretical applicability with respect to the events which transpire in the physical world. Notice that it is "single case *application*", not single case *definition*, which Mellor demands at the close of the paragraph quoted above. Mellor does not criticize the account as such, but complains that I prefer the term "weight" to the term "probability" in this context. However, if a theory tells us how to derive an appropriate value of the weight from knowledge of probabilities, and tells us how to use these probabilistically-based weights for making appropriate decisions regarding single cases, one cannot legitimately complain that the theory fails to make sense of the treatment of the single case. There may be valid grounds for criticism of the solution which is offered, but that is a far cry from suggesting that no serious attempt at a solution was ever proffered.

Mellor also complains that I "relegate" ascertainment of single case weights to the realms of methodology rather than to probability theory proper. In drawing this distinction (1967, pp. 92–93), I made an explicit comparison with Carnap's distinction between inductive logic proper (which deals with the logical relations upon which degrees of confirmation are based) and the methodology of induction (which includes such rules as the requirement of total evidence and the rule of maximizing estimated utility). One can hardly claim that methodology is, for Carnap, a wastebasket for practical problems which have little philosophical interest or importance. One might as well say that principles of measurement – such as paths of light rays are null geodesics, while freely falling gravitational test particles follow non-null geodesics – are *mere* principles of the application of geometrical theory to the physical world. Recall once more that Mellor's demand is for "single case *application*"; Carnap explicitly characterized methodology in terms of problems of application (1950, §44), and I certainly had that conception in mind when discussing the problem of the single case. Mellor thus disparages my approach for attempting to do exactly what he regards as indispensable.

I am in total agreement with Mellor when he says, "The issue cannot be disposed of *just* by *calling* single case chances 'weights' and the problems of determining their values 'methodological' (p. 54, my italics). Frequentists have not merely introduced new *terminology* (as Mellor's remarks suggest); they have articulated a theory of application of probability to the single case (which Mellor does not deign to discuss). Moreover, frequentists have been sensitive to the difficulties in their theory. Reichenbach points out that, in practice, his approach may not yield unique values, and he indicates how practical constraints may dictate greater emphasis upon narrowing the reference class (thus reducing the body of statistical evidence) or upon increasing the available data (thus employing a somewhat broader reference class). When application of knowledge to practical situations is at issue, it is not surprising that practical considerations may loom large.

In my approach to the problem of the single case, I have made use of the concept of a "homogeneous reference class". It has long been evident that a precise characterization of homogeneity poses severe problems. In a recent paper (1977) I have systematically attacked them. I am sure that this account is not totally satisfactory, but the problems do not seem insuperable. It may be that any given single case can be classified in only one maximal homogeneous reference class, thus furnishing a unique single case weight, but I do not know of any argument to demonstrate that conclusion convincingly. Hence, uniqueness remains a problem for frequentists in their treatment of the single case.[12]

Consider the following example. A particular die is tossed repeatedly on a particular table, with the result that the long run frequency of side 6 showing converges to 1/6. Closer examination of the sequence of tosses S reveals, however, that in the subsequence S_e consisting of even-numbered members of S, the limiting frequency of 6 has a different value, say 1/4, while in the subsequence S_o consisting of odd-numbered members of S, the limiting frequency of 6 is 1/12. Within the even-numbered subsequence S_e, let us suppose, the outcome 6 occurs randomly; there is no place selection which will pick out a subsequence of S_e in which the limiting value of the frequency of 6 is different from 1/4. We find, furthermore, that within the prime-numbered subsequence S_p of S the limiting frequency of 6 has still a different value, say 1/96. The question we now raise is what weight should be attached to the outcome 6 on the second toss. Since 2 is an even number, we are tempted by the value 1/4; but since it is also prime, the value 1/96 is appealing. Because the class of even primes is finite – indeed, it is a unit class – there is no well-defined limiting frequency for the intersection of S_e and S_p. We reject the alternative which says that the probability within this unit class is 1 or 0 unless we can refer it to a hypothetical sequence in which the limiting frequency has one or the other of these values.

We are *not* prepared to believe that the foregoing statistical regularities have no physical explanation; therefore, we examine the physical mechanisms involved. Concealed within the die we find a small but powerful magnet, and under the table we find an electromagnet whose polarity can can be reversed at the touch of a hidden button. Upon further experimental testing, or on the basis of theoretical analysis of the physical set-up, we find that the reversal of polarity between successive tosses of the die suffices to explain quite adequately the limiting frequencies of 1/4 in S_e and 1/12 in S_o. The propensity theorist says that these physical features of the gaming device produce propensities which are different from those found in unbiased die-tossing devices. The frequentist says that tosses of an unbiased die, tosses of the magnetized die with a particular orientation of the field of the electromagnet, and tosses of the magnetized die with the opposite orientation of the field of the electromagnet constitute three distinct reference classes, and each has a different probability (i.e., limiting frequency) for 6 to show. Notice that the frequentist describes his reference classes in terms of the physical circumstances under which the trials are made – in just about the same way as the propensity theorist characterizes his propensities.

We have not yet explained the low frequency of 6 among the prime tosses. A further search reveals a second electromagnet, much more powerful than the first, also concealed beneath the table. It is turned on during each prime toss; otherwise it is off. We have no trouble in determining that the field of the second magnet, superimposed upon that of the first magnet with the polarity of the odd tosses, explains the limiting frequency of 1/96 in the sequence consisting of odd-prime tosses; again the frequency and propensity theorists are in substantial agreement. Since we are talking about long run frequencies, we can add the second toss to the sequence of odd-prime tosses, with the result that the limiting frequency in the sequence of prime tosses is the same as the limiting frequency in the sequence of odd-prime tosses. But what about toss number two? Is the frequentist compelled to say that its probability or weight is 1/96? Mellor would apparently have us believe so. He quotes L. J. Savage in the claim that frequentists "hold that...evidence...for the magnitude of the probability...is to be obtained by observations of some repetitions of the event, and from no other source whatever" (p. 50).

The propensity theorist is in no such bind. Since "propensity shares with the other dispositions of a chance set-up a subjection to the principle of connectivity", the propensity is associated with other physical features of the system. We must look at trial number two to see whether the second magnet is turned on, and if so, whether its polarity is the same or opposite, so that it reinforces the field of the first magnet. In that case, the propensity for 6 on the second trial would be very high indeed. Its value could be determined theoretically, or by repeated trials. If the propensity theory stimulates us to look for relevant physical features of the chance mechanism, that is all to the good.

At the same time, it is evident from Reichenbach's discussion, quoted above, of the case in which one result in a sequence is obtained by placing the die on the table with face 6 uppermost, that he would follow the same procedure as the propensity theorist. His analysis would, of course, bear a strong resemblance to the "virtual sequence propensity interpretation", for he would say that, in a long sequence of trials conducted under just these conditions, the limiting frequency of 6 would be (say) 95/96. Although there is a sense in which Reichenbach maintains that knowledge of probabilities rests *ultimately* upon frequency data, there is no sense in which he can be correctly charged with the view that counting frequencies is the *only* source of information about probabilities. Sections 69–70 of his (1949) explicitly discuss *various* ways of establishing knowledge of probability values. It is possible, of course, that we might not have been as lucky as supposed in the foregoing example in finding the *relevant* physical characteristics of the mechanism which was operative in toss number two, but the intractable cases will pose as much difficulty for the propensity theorist as for the frequentist. If such problems arise, we may have to admit that, for all we know, there is no objectively correct unique value for the single case probability.

As I indicated above, in the comparison of Popper's and Mellor's version of the propensity theory, uniqueness is a fundamental problem. Mellor's way of dealing with it led him to the conclusion that chance does not exist if determinism is true. On Popper's approach, as nearly as I can tell, probabilities are relational with respect of certain features of the physical set-up; thus, trial number two could have one propensity relative to some features of the mechanism and another propensity relative to other features. Perhaps it is better to allow that single case probabilities need not be unique and non-relational.

Regardless of our success or failure in defining objective homogeneity, in most practical situations we do not have enough knowledge to find a fully homogeneous reference class, and to assign the probability within that class to the single case in question. Thus, we face just the kinds of issues Reichenbach raised regarding practical ambiguities in such situations. Nevertheless, roughly speaking, we adopt something quite analogous to Carnap's requirement of total evidence – that is, we characterize our reference class in terms of all known statistically relevant factors. "The frequency analogue of such a 'total evidence' requirement", says Mellor (p. 53), "is that the chance that a is G should be identified with the frequency of G in the most closely defined class of which a is a member, namely the class of things that are F_1 *and* F_2 *and* ... F_i *and* ...; but unless some limit is set to the increasingly detailed specification of this 'reference class', a will be its sole member". It is precisely the consideration of objections of this sort which led me to reformulate Reichenbach's rule of selecting *the narrowest class for which reliable statistics are available* as the rule of selecting the *broadest homogeneous reference class*. Taking the latter formulation (which, I believe, expresses Reichenbach's intent), we see that the effort is not to keep narrowing the reference class indefinitely, but rather, keeping it as broad as possible – provided, of course, that we partition it in terms of statistically relevant properties F. It is not evident *a priori* that we will end up with nothing but unit reference classes. Given Avogadro's number of radioactive nuclei of the same isotope – say C^{14} – there is no reason of which I am aware for supposing that this reference class can be relevantly subdivided with respect to spontaneous radioactive decay. This sort of example is one of Mellor's three paradigms (p. xi). Indeed, it seems to me that the very concept of indeterminism can be characterized quite appropriately by saying that it implies the existence of objectively homogeneous reference classes in which the probability of the occurrence is neither 1 nor 0. For Mellor, chance does not exist unless there are, in fact, reference classes of this sort.

A further point should be added to this discussion on the frequency theory of the single case. In a (1957) paper, A. J. Ayer raised what seems to me to be a profound difficulty concerning Carnap's requirement of total evidence. To the best of my knowledge, Carnap never answered it; indeed, at an International Philosophy Congress held in Mexico City in 1963 which I attended, Ayer publicly posed the problem to Carnap, and Carnap did not seem even to appreciate its significance. Be that as it may, Ayer goes on to argue that essentially the same problem arises for the frequentist in dealing with the problem of the single case. In my (1967, pp. 95–96) I argued that the frequentist *has* an answer to that problem; this solution was given by Reichenbach (1949, §72, 56) and is based upon a carefully developed argument to the effect that persistent use of the frequentist strategy will pay off in the long run. No analogous argument is available to the logical theorist. These considerations are dismissed by Mellor with the two words, "*pace* Salmon" (p. 53). Since the problem of probability of single cases is *the* pivotal issue between propensity theorists and frequency theorists, it seems to me that the propensity theorists owe us a deeper consideration of the frequentist approach than Mellor has furnished.

5. HUMEAN QUALMS

Mellor acknowledges from the outset the "need to show that no properly hallowed Humean doctrine is denied" (p. 3), and in the final chapter he returns to this issue. He immediately assures us that "Chance is not a sort of weak or intermittently successful causal link" (p. 151). On Mellor's account, in contrast to Popper's, it will be recalled, a propensity is a *disposition* to display a chance distribution on every trial, rather than a *tendency* to produce a particular result (e.g., heads) in a certain percentage of trials. On his account, therefore, "a statistical law asserts of *each* trial of a certain kind that on it there is the stated chance p of some outcome" (p. 151). If a Humean is prepared to sanction universal laws of the form

(1) All F are G

he should have no objection to those of the form

(2) All F have a chance p of being G

which is just as universal and just as lawful as (1). For an F to *have a chance p of being G* is, of course, for F to have a certain disposition, but that should not occasion misgivings on the part of the Humean, for surely the "G" in (1) must be allowed to stand for dispositional properties (e.g., solubility). If good sense can be made of the notion of displaying a chance distribution (rather than a definite outcome) on a single trial, then I think Mellor's claim to avoid Humean pitfalls is well-founded. As I have argued above, however, it does not seem to me that he has adequately explicated this concept. For Mellor, "The chance of the coin falling heads when tossed is. . . the measure of that reasonable partial belief [in the outcome heads]" (p. 2). I do not believe Mellor has provided a satisfactory account of reasonable partial belief.

It is not my intention to argue that no account of reasonable partial belief can be constructed. Suppose, then, that one has been provided, or that I am wrong in claiming that Mellor's account is unsatisfactory. Even granting all this, it seems to me, Mellor pays a high price to remain in the camp of the Humeans. In his attempt to show that no Humean strictures are violated, he shows that his account of rationality hinges solely upon the frequencies with which certain outcomes occur in particular sorts of repeated trials – or would occur if many trials of these sorts were made. Recall the crucial role, in his analysis, of the law of large numbers, and of its affirmation concerning the relationship between the betting quotient and the frequency. I fail to find any point in the argument at which any essential use is made of the chance of a given outcome on a *single* trial. The "cash value" of the theory seems extremely close, if not equal, to that of the frequency theory.

The point can be illustrated by means of Laplace's famous case of the biased coin, to which Mellor refers in several places (pp. 129–136, 165–167). The situation is this. A coin, which is about to be tossed for the first time, is known to be biased, but it is unknown whether the bias favors heads or tails. Laplace, appealing to the principle of indifference, argued that the probability of heads on the first toss is 1/2. Mellor's analysis of this example concludes that two distinct propensities are involved. Let us consider the physical details of the situation more closely. Suppose that a machine stamps out disks, which will be embossed by a different machine to make them into coins. The disks stamped out by the first machine have some physical characteristic which biases one side to come up much more often than the other if the disks are flipped repeatedly in the standard manner. However, at this stage, we cannot say that a given disk has a bias for heads or for tails, for the two faces of the disk have not yet been embossed with the insignia known as heads and tails. Assume further that the process of embossing, as carried out by the second machine, does nothing to alter the bias (it is always assumed that the markings which distinguish the two sides of a coin have no effect on its chance of landing with a particular side uppermost), and that the embossing machine randomly marks the coins in such a way that half of them are embossed with heads on the side which is more likely to show, while the other half are embossed with tails on that favored side. Thus, for the sake of definiteness, let us say that each given coin has either (1) the propensity to display the chance distribution {chance of heads = 1/4; chance of tails = 3/4} or (2) the propensity to display the chance distribution {chance of heads = 3/4; chance of tails = 1/4}. The embossing machine also has a propensity to display the following chance distribution: {chance of favored side being marked heads = 1/2; chance of favored side being marked tails = 1/2}, Now, when a given coin is picked randomly from the output of the minting machines and tossed, what propensity is being displayed? Mellor's answer is,

> Two partial beliefs seem to be reasonable on the first toss of a biased coin: the CBQ 1/2 displaying a propensity of the labelling set-up, and some other CBQ displaying the unknown bias of the coin. Which is really the more reasonable depends, on the above account, on what counts as repeating the bet. If further biasing of labelled coins, or labellings of biased coins, are what the gambler must break even on, his degree of partial belief should be 1/2; if further tosses of the same biased coin, it should be something other than 1/2. (p. 165)

In other words, whether the chance of heads on this toss is 1/2 or some other value depends upon whether further trials will consist of additional throws of the same coin or of initial tosses of other coins produced in the manner described.

I cannot see what is gained in adding to the foregoing analysis that on *this particular* first toss of *this particular* coin there is a unique, objective, non-relational chance of heads, if identification of this chance depends "on what counts as repeating the bet". It seems more straightforward to say this particular event belongs to two possible sequences, and the relative frequency of heads in the one differs from its relative frequency in the other. Whether a gambler "breaks even" in Mellor's sense depends upon his betting quotient and upon which sequence he plays.

In his introductory apology for his book, Mellor acknowledges that "[t]he ingredients of the present theory are in the literature, but they have hitherto been no more than half baked". Under Mellor's culinary art, it seems to me, they come out of the oven a bit overdone. In particular, his thesis that propensities are dispositions to display chance distributions seems less palatable than Popper's contention that propensities are tendencies to yield some specified outcome in a given percentage of cases.

In this more standard form, the propensity theory does encounter Humean objections – propensities are more like "weak or intermittently successful causal links".[13] If a glass is fragile, dropping it on a hard surface *causes* it to break. If a coin is unbiased, then tossing it in the standard manner *produces* heads in 1/2 of the trials. On Hume's analysis of causation, causal relations involve constant conjunctions; the meaning of "F causes G" involves the general law, "Every F is followed by a corresponding G". If we want to assert that this particular instance f_1 caused that particular instance g_1, we are on perfectly solid ground, provided we can derive it from the foregoing general law. To establish the particular causal relationship between the two individuals f_1 and g_1 without reference to the general law is proscribed. In dealing with the probabilities (or weights) to be assigned to single events, the frequentist follows the Humean tack. Between the *classes* A and B, he says, there is a certain probability relation – the long run frequency with which members of A also belong to B. It is perfectly legitimate to appeal to this general statistical relationship (with appropriate attention to such matters as the homogeneity of the reference class A) to say that there is probability p that a particular individual $a_1 \in A$ is a B. If, however, one attempts to develop a conception of probability according to which probabilities are assigned fundamentally to single events, without being derived from relations between classes, this constitutes an attempt to do, in the realm of statistical law, exactly what the Humean analysis forbids in the realm of universal causal law.

Whether the Humean analysis of causal laws and individual causal relations is correct is, of course, a matter of considerable controversy. I do not wish to argue the point here. It is important to realize, however, that many philosophers have argued with some force and vivacity that it is neither desirable nor possible to develop an account of individual causal relations which are logically prior to general causal laws. A similar argument can be made in the realm of statistical laws. Perhaps it is neither desirable nor possible to develop a single case interpretation of probability which does not depend logically upon relative frequencies within classes of occurrences. It may be intuitively evident to some (e.g., Mellor) that the frequentist has put the cart before the horse, but in the realm of probability – perhaps even more than in other areas of philosophical investigation – fundamental intuitions should be subjected to close scrutiny. They may turn out to be false!

6. CONCLUDING REFLECTIONS

In spite of the foregoing criticisms, I am inclined to conjecture that the concept of propensity *can* be used in a way which will contribute fruitfully to our understanding of the world. As I indicated clearly in §2 above, I do not accept Popper's claim that the propensity interpretation of probability can take the mystery out of quantum theory. Nor can I accept Mellor's analysis of propensities as dispositions to display chance distributions. I would suggest, instead, that they may be taken to represent just the sort of weak intermittent causal link which Mellor emphatically eschews. In construing propensities in this fashion, I would want to deny that they provide an interpretation of the probability calculus. Instead, I would stick to the view that the frequency interpretation is the most appropriate one – at least in contexts in which we are dealing with statistical laws and physical probabilities.

In a recent article (1977a), I have tried to show how certain spatiotemporally continuous physical processes can be considered as means for the propagation of causal influence. In many of the cases in which we talk about cause-effect relations between distinct events, there is a physical connection between the cause and the effect which consists precisely in just such a continuous causal process. Such physical processes occur in the body, for example, when a tap on the knee causes the foot to jerk. These cause-effect relations need not be deterministic; it seems to me that striking a golf ball in the direction of a pane of glass may cause the glass to shatter even though we need not assume that breakage invariably occurs in such circumstances. In such cases, I believe, it is legitimate to say that the process constituted by the moving golf ball transmits a propensity to shatter panes of glass of given specifications – a disposition which can be realized only if a pane of glass of the right kind happens to lie in the path of the ball. In a situation of this sort, it might be appropriate to identify the strength of the propensity with the measure of the probability of the result.

If we construe the cause-effect relation in the manner here suggested, there is no difficulty whatever in talking about singular causal relations between individual events rather than classes of events. Event *a* can be said to cause event *b* (where "*a*" and "*b*" stand for particular events) if there is a causal process of the appropriate sort joining them. This type of statement is legitimate whether the causal connection is deterministic or probabilistic. Propensities can thus apply unambiguously to individual cases, but they will be relational.

Popper has compared the frequency interpretation with the propensity interpretation in the following terms:

> The frequency interpretation always takes probability as relative to a sequence which is assumed as given; and it works on the assumption that a probability is *a property of some given sequence*. But with our modification, the sequence in its turn is defined by its set of *generating conditions*; and in such a way that probability may now be said to be *a property of the generating conditions* (1960, p. 34, Popper's italics).

There is another – and I think better – way of characterizing the situation. Going back to Reichenbach's version of the frequency theory, as described above, we must say that a *pair* of sequences is assumed, and that the probability refers to a relation between the two sequences. Thus, for example, we have a sequence of tosses of a given die, and a sequence of corresponding instances of the die coming to rest on the table. The probability of side 6 showing is the relative frequency with which members of the first sequence (tosses) are paired with members of the second sequence (landings) having the property that the side 6 is uppermost. In examples of this sort, a member of the first sequence is an event which initiates a causal process – the tumbling of the die – which leads to a certain outcome in a certain percentage of cases. I suggest, in examples of this sort, retaining the frequency interpretation of probability, and using the concept of *propensity* to characterize – not the "generating conditions" as Popper would have it – but rather, the strength of the tendency of the connecting (probabilistic) causal process to produce the outcome in question. Hume made us chary of talking about causal connections, but it seems to me that we *can* legitimately identify certain physical processes as causal connections.

As Paul W. Humphreys has pointed out in a private communication, there is an important limitation upon identifying propensities with probabilities, for we do not seem to have propensities to match up with "inverse" probabilities. Given suitable "direct" probabilities we can, for example, use Bayes's theorem to compute the probability of a particular cause of death. Suppose we are given a set of probabilities from which we can deduce that the probability that a certain person died as a result of being shot through the head is 3/4. It would be strange, under these circumstances, to say that this corpse has a propensity (tendency?) of 3/4 to have had its skull perforated by a bullet. Propensity can, I think, be a useful causal concept in the context of a probabilistic theory of causation, but if it is used in that way, it seems to inherit the temporal asymmetry of the causal relation.

Consider another example of a type propensity theorists tend not to bring up. Suppose that a company which manufactures light bulbs has two factories. Factory *A* is somewhat older than Factory *B*; the equipment in Factory *B* is newer, more advanced, and more efficient. Factory *B* produces a larger number of bulbs, and those which come from Factory *B* tend to last longer before burning out. Suppose, further, that the company guarantees its bulbs to burn for at least 1500 hours. A particular light bulb burns out before the allotted time; we ask for the probability that it was produced by Factory *B*. This is a perfectly reasonable question about probabilities, and with numerical values for the rates at which the two factories produce light bulbs and the relative frequencies with which they produce defective bulbs, it is easy to find the answer to the question. However, it would seem strange to speak of the premature failure of the light bulb as a "generating condition" of its having been manufactured in Factory *B*. I would find it odd to speak of the propensity of the bulb to have been produced by the factory in question, but quite straightforward to speak of the propensity of the factory to contribute to the total output, and of the propensity of each factory to produce faulty bulbs. As I said above, the notion of propensity seems appropriate when we can find causal processes to which to apply it.

Propensities, considered in this way, may play a useful role in quantum theory, where it seems natural to identify wave amplitudes with propensities – in many cases, at least, with propensities to actuate various kinds of measuring apparatus. In this context, of course, wave amplitudes must not be identified with probabilities, so propensities could not be probabilities. However, since the careful elaboration of this proposal involves many technical difficulties, I shall not attempt it here. If this tentative suggestion were to be developed successfully, it might help us to achieve a better understanding of the microphysical domain, but I do *not* believe it could take the mystery out of quantum theory.

I hope it is evident that, although I have expressed sharp disagreement with much that Mellor has said, I have found his book extremely provocative and useful. With verve and style, it brings clearly to the fore a wide range of the most significant issues which have arisen in the philosophy of probability in the last quarter century. That is a valuable contribution.

University of Arizona

NOTES

* The author wishes to express his gratitude to the National Science Foundation (U.S.A.) for support of research.

[1] A bench mark is "a surveyor's mark made on a permanent landmark that has a known position and altitude; bench marks are used as reference points in determining other altitudes..." *Webster's New World Dictionary of the American Language*, World Publishing Company, Cleveland, 1966. Kyburg (1964) uses this apt expression for the same works in his survey of inductive logic.

[2] Ian Hacking (1975) has argued persuasively that the concept of probability could not emerge until the notions of chance and degree of partial evidential support were both available and were brought into contact with one another.

[3] F. P. Ramsey and Bruno de Finetti were earlier contributors to this approach, but Savage (1954) was the work which captured the interest of modern philosophers and statisticians.

[4] The same general program of strengthening personalism is pursued by Abner Shimony (1970) in his elaboration of "tempered personalism".

[5] The usage being introduced here differs from Carnap's.

[6] See Salmon (1965) for a more detailed analysis. This analysis assumes that we are dealing with a Bernoulli sequence of trials, an assumption which requires empirical support in any concrete situation.

[7] For Giere (1973), who does not try to base his propensity theory on rational degree of belief, there is no problem about the second level probability; it is simply another propensity.

[8] Giere (1973) explains this distinction clearly, and provides further documentation of this claim about the ambiguity in Popper's writings. In contrast to Mellor and Giere, Hacking (1965) explicitly adopts a virtual sequence propensity theory.

[9] Giere (1973), in contrast to Mellor, explicitly requires his single-case propensities to be probabilities. In contrast to Popper, Giere unequivocally adopts a single-case theory.

[10] As a result of an unfortunate misprint, the word "tendency" was omitted from this passage, as indicated by the square brackets.

[11] I should emphasize that almost all of my recent work on single case probability or weight has been motivated by considerations in the theory of scientific explanation; see, for example, Salmon, *et al.* (1971), §4.

[12] Although Benenson (1977) considers the problem of uniqueness at length, I do not believe he offers an adequate solution. It seems to me that his treatment does not handle the example of the probability of 6 on the second toss of the magnetic die.

[13] Giere (1973, p. 478) acknowledges explicitly the non-Humean character of his propensities.

REFERENCES

Ayer, A. J., 1957, 'The Conception of Probability as a Logical Relation' in S. Körner (ed.), *Observation and Interpretation*, Butterworths Scientific Publications, London.

Benenson, F., 1977, 'Randomness and the Frequency Definition of Probability', *Synthese*, **36**, 207–233.

Carnap, R., 1950, *Logical Foundations of Probability*, University of Chicago Press, Chicago.

Carnap, R., 1962, 'The Aim of Inductive Logic' in E. Nagel, *et al.* (eds.), *Logic, Methodology and Philosophy of Science*, Stanford University Press, Stanford.

Fetzer, J., 1974, 'Statistical Probabilities: Single Case Propensities vs. Long Run Frequencies' in W. Leinfellner and E. Köhler (eds.), *Developments in the Methodology of Social Science*, D. Reidel Publishing Co., Dordrecht, pp. 387–397.

Giere, R., 1973, 'Objective Single-Case Probabilities and the Foundations of Statistics' in P. Suppes *et al.* (eds.), *Logic, Methodology and Philosophy of Science*, North-Holland, Amsterdam, pp. 467–483.

Hacking, I., 1965, *Logic of Statistical Inference*, The University Press, Cambridge.

Hacking, I., 1975, *The Emergence of Probability*, Cambridge University Press, Cambridge.

Kyburg, H., 1964, 'Recent Work in Inductive Logic', *American Philosophical Quarterly* **1**, 249–287.

Kyburg, H., 1974, 'Propensities and Probabilities', *British Journal for the Philosophy of Science* **25**, 358–375.

Popper, K., 1957, 'The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory' in S. Körner (ed.), *Observation and Interpretation*, Butterworths Scientific Publications, London, pp. 65–70.

Popper, K., 1960, 'The Propensity Interpretation of Probability,' *British Journal for the Philosophy of Science* **10**, 25–42.

Popper, K., 1967, 'Quantum Mechanics without "The Observer"' in M. Bunge (ed.), *Quantum Theory and Reality*, Springer-Verlag, New York, pp. 7–44.

Ramsey, F. P., 1931, 'Truth and Probability' in R. Braithwaite (ed.), *The Foundations of Mathematics and other Logical Essays*, Routledge and Kegan Paul, London.

Reichenbach, H., 1949, *The Theory of Probability*, University of California Press, Berkeley & Los Angeles.

Salmon, W., 1965, 'What Happens in the Long Run', *Philosophical Review* **74**, 373–378.

Salmon, W., 1967, *The Foundations of Scientific Inference*, University of Pittsburgh Press, Pittsburgh.

Salmon, W., *et al.*, 1971, *Statistical Explanation and Statistical Relevance*, University of Pittsburgh Press, Pittsburgh.

Salmon, W., 1977, 'Objectively Homogeneous Reference Classes', *Synthese* **36**, 399–414.

Salmon, W., 1977a, 'An "At-At" Theory of Causal Influence', *Philosophy of Science* **44**, 215–224.

Savage, L. J., 1954, *Foundations of Statistics*, John Wiley & Sons, New York.

Schilpp, P. (ed.), 1963, *The Philosophy of Rudolf Carnap*, Open Court, La Salle, Ill.

Shimony, A., 1970, 'Scientific Inference' in R. Colodny (ed.), *The Nature and Function of Scientific Theories*, University of Pittsburgh Press, Pittsburgh.

Manuscript received 16 November 1978