Postscripts to "A Subjectivist's Guide to Objective Chance"

A. NO ASSISTANCE NEEDED[1]

Henry Kyburg doubts that the Principal Principle has as much scope as my praise of it would suggest. He offers a continuation of my questionnaire, says that his added questions fall outside the scope of the Principal Principle, and suggests that we need some Assistant Principle to deal with them. His first added question is as follows.[2]

> *Question.* You are sure that a certain coin is fair. It was tossed this morning, but you have no information about the outcome of the toss. To what degree should you believe the proposition that it landed heads?
>
> *Answer.* 50 per cent, of course.

That's the right answer (provided the question is suitably interpreted). But the Principal Principle, unassisted, does suffice to yield that answer. What we must bear in mind is that the Principle relates time-dependent chance to time-dependent admissibility of evidence; and that it applies to any time, not only the present.

Kyburg thinks the Principle falls silent "since there is *no* chance that the coin fell other than the way it did," and quotes me to the effect that "what's past is no longer chancy." Right. We won't get anywhere if we apply the Principle to *present* chances. But what's past *was* chancy, if indeed the coin was fair; so let's see what we get by applying the Principle to a past time, and working back to present credences. Notation:

[1] In writing this postscript, I have benefited from a discussion by W. N. Reinhardt (personal communication, 1982). Reinhardt's treatment and mine agree on most but not all points.

[2] Henry E. Kyburg, Jr., "Principle Investigation," *Journal of Philosophy* 78 (1981): 772–78.

t: a time just before the toss,
C: a reasonable initial credence function that will yield my later credences by conditionalizing on total evidence,
C_0: my present credence function,
A: the proposition that the coin fell heads,
X: the proposition that the coin was fair, that is that its chance at t of falling heads was 50%,
E: the part of my present total evidence that is admissible at t,
F: the rest of my present total evidence.

Since *ex hypothesi* I'm certain of X, we have

(1) $C_0 = C_0(-/X)$.

By definition of C, we have

(2) $C_0 = C(-/EF)$.

Assuming that F is irrelevant to the tosses, we have

(3) $C(A/XEF) = C(A/XE)$.

By the Principal Principle, applied not to the present but to t, we have

(4) $C(A/XE) = 50\%$.

Now, by routine calculation from (1)–(4) we have

(5) $C_0(A) = 50\%$.

which answers Kyburg's question.

Step (3) deserves further examination, lest you suspect it of concealing an Assistant Principle. Recall that F is the part of my present total evidence that was not admissible already at time t. Presumably it consists of historical information about the interval between t and the present. For historical information about earlier times would be already admissible at t; and historical information about later times, or nonhistorical information, could scarely be part of my present total evidence. (Here, as in the paper, I set aside strange possibilities in which the normal asymmetries of time break down. So far as I can tell, Kyburg is content to join me in so doing.) Thus if I had watched the toss, or otherwise received information about its outcome, that information would be included in F.

However, Kyburg stipulated in his question that "you have no information about the outcome of the toss". We might reasonably construe that to mean that no information received between t and the present is evidentially relevant to whether the coin fell heads, with evidential relevance construed in the usual way in terms of credence. Then (3) comes out as a stipulated condition of the problem, not some extra principle.

There is a different, stricter way that Kyburg's stipulation might perhaps be construed. It might only exclude information that settles the outcome decisively, leaving it open that I have information that bears evidentially on the outcome without settling it. For instance, it might be that the tosser promised to phone me if the toss fell heads, I got no phone call, but that is far from decisive because my phone is not reliable. On that construal, we are not entitled to assume (3). But on that construal Kyburg's answer is wrong; or anyway it isn't right as a matter of course on the basis of what he tells us; so we don't want any principle that delivers that answer.

Kyburg has a second added question to challenge the Principal Principle.

> *Question.* As above, but you know that the coin was tossed 100 times, and landed heads 86 times. To what degree should you believe the proposition that it landed heads on the first toss?
>
> *Answer.* 86 per cent.

The strategy for getting the Principal Principle to yield an answer is the same as before, but the calculation is more complicated. Notation as before, except for

A: the proposition that the coin fell heads *on the first toss*,

B: the proposition that the coin fell heads 86 times out of 100,

X: the proposition that the coin was fair, that is that its chance at t of falling heads was 50% *on each toss*,

F: the rest of my present total evidence, besides the part that was admissible at t, *and also besides the part B*,

x: the fraction of heads-tails sequences of length 100 in which there are 86 heads.

Our equations this time are as follows. They are justified in much the same way as the like-numbered equations above. But this time, to get the new (2) we split the present total evidence into three parts B, E, and F. And to get the new (4), we use the Principal Principle repeatedly to multiply endpoint chances, as was explained in the section of the paper dealing with chance of frequency.

(1) As before;

(2) $C_0 = C(-/BEF)$;

(3) $C(A/XBEF) = C(A/XBE)$;

(4) $C(AB/XE) = x{\cdot}86\%$, $C(B/XE) = x$;

(5) $C_0(A) = 86\%$.

Kyburg also thinks I need an extra "Principle of Integration" which I neglected to state. But this principle, it turns out, has nothing especially to do with chance! It is just a special case of a principle of infinite additivity for credences. Indeed it could be replaced, at the point where he claims I tacitly used it, by *finite* additivity of credences. (And finite additivity goes without saying, though I nevertheless did say it.) To be sure, if we want to treat credences in the setting of nonstandard analysis, we are going to want some kind of infinite additivity. And some kind of infinite additivity comes automatically when we start with finite additivity and then treat some infinite sets as if they were finite. It is an interesting question what kind of infinite addivity of credences we can reasonably assume in the nonstandard setting. But this question belongs entirely to the theory of credence—not to the connection between chance and credence that was the subject of my paper.

B. CHANCE WITHOUT CHANCE?

Isaac Levi thinks that I have avoided confronting "the most important problem about chance"; which problem, it seems, is the reconciliation of chances with determinism, or of chances with different chances.[3] Consider a toss of coin. Levi writes that

> . . . in typical cases, the agent will and should be convinced that information exists (though inaccessible to him) which is highly relevant [to the outcome]. Thus, the agent may well be convinced that a complete history through [the onset of the toss] will include a specification of the initial mechanical state of the coin upon being tossed and boundary conditions which, taken together, determine the outcome to be heads up or tails up according to physical laws.
>
> . . . given the available knowledge of physics, we cannot [deny that the mechanical state of the coin at the onset of the toss determines the out-

[3] Isaac Levi, review of *Studies in Inductive Logic and Probability*, ed. by R. C. Jeffrey, *Philosophical Review* 92 (1983): 120–21.

> come] provided we can assume the motion of the coin . . . to be sealed off from substantial external influences. But even if we allow for fluctuations in the boundary conditions, we would not suppose them so dramatic as to permit large deviations from 0 or 1 to be values of the chances of heads. . . .

And yet

> Lewis, however, appears ready to assign .5 to the chance of [the] coin landing heads up. . . .

So how do I square the supposition that the chance of heads is 50% with the fact that it is zero or one, or anyway it does not deviate much from zero or one?

I don't. If the chance is zero or one, or close to zero or one, then it cannot also be 50%. To the question how chance can be reconciled with determinism, or to the question how disparate chances can be reconciled with one another, my answer is: *it can't be done.*

It was not I, but the hypothetical "you" in my example, who appeared ready to assign a 50% chance of heads. If my example concerned the beliefs of an ignoramus, it is none the worse for that.

I myself am in a more complicated position than the character in this example. (That is why I made an example of him, not me.) I would not give much credence to the proposition that the coin has a chance of heads of 50% exactly. I would give a small share of credence to the proposition that it is zero exactly, and an equal small share to the proposition that it is one exactly. I would divide most of the rest of my credence between the vicinity of 50%, the vicinity of zero, and the vicinity of one.

The small credence I give to the extremes, zero and one exactly, reflects my slight uncertainty about whether the world is chancy at all. Accepted theory says it is, of course; but accepted theory is not in the best of foundational health, and the sick spot (reduction of the wave function brought on by measurement) is the very spot where the theory goes indeterministic. But most of my credence goes to the orthodox view that there are plenty of chance processes in microphysics. And not just the microphysics of extraordinary goings-on in particle accelerators! No; for instance the making and breaking of chemical bonds is chancy, so is the coherence of solids that stick together by means of chemical bonding, so is the elasticity of collisions between things that might bond briefly before they rebound,. . . . So is any process whatever that could be disrupted by chance happenings nearby—and infallible "sealing off" is not to be found.

In Levi's physics, a coin coming loose from fingers and tumbling in

air until it falls flat on a table is a classical system, an oasis of determinism in a chancy microworld. I do not see how that can be. The coin, and the fingers and the air and the table, are too much a part of that microworld. There are also the external influences, which cannot be dismissed either by requiring them to be substantial or by invoking fictitious seals; but never mind, let us concentrate on the toss itself. There is chance enough in the processes by which the coin leaves the fingers; in the processes whereby it bounces off air molecules and sends them recoiling off, perhaps to knock other molecules into its path; in the process whereby the coin does or doesn't stretch a bit as it spins, thereby affecting its moment of inertia; and in the processes whereby it settles down after first touching the table. In ever so many minute ways, what happens to the coin is a matter of chance.

But all those chance effects are *so* minute.—But a tossed coin is *so* sensitive to minute differences. Which dominates—minuteness or sensitivity? That is a question to be settled not by asking what a philosopher would find it reasonable to suppose, but by calculation. The calculations would be difficult. We may not make them easier by approximations in which expected values replace chance distributions. I have not heard of anyone who has attempted these calculations, and of course they are far beyond my own power. Maybe they are beyond the state of the art altogether. Without them, I haven't a clue whether the minuteness of the chance effects dominates, in which case the chance of heads is indeed close to zero or one; or whether instead the sensitivity dominates, in which case the chance of heads is close to 50%. Hence my own distribution of credence.

The hypothetical "you" in my example has a different, simpler distribution. Why? He might be someone who has done the calculations and found that the sensitivity dominates. Or he might have been so foolish as to intuit that the sensitivity would dominate. Or he might be altogether misinformed.

Well-informed people often say that ordinary gambling devices are deterministic systems. Why? Perhaps it is a hangover of instrumentalism. If we spoke as instrumentalists, we would be right to say so—meaning thereby not that they really *are* deterministic, but rather that it is sometimes instrumentally useful to pretend that they are. To the extent that it is feasible to predict gambling devices at all—we can't predict heads or tails, but we can predict, for instance, that the coin won't tumble in mid-air until next year, and won't end up sticking to the wall—deterministic theories are as good predictive instruments as can be had. Perhaps when the instrumentalist expert says that tossed coins are deterministic, the philosopher misunderstands him, and thinks he means that tossed coins are deterministic.

Can it be that Levi himself was speaking as an instrumentalist in the passages I cited? If so, then the problem of reconciling chance and determinism is not very hard. It is just the problem of reconciling truth *simpliciter* with truth in fiction. In truth, nobody lived at 221B Baker Street; in fiction, Holmes lived there. In truth, most likely, the coin is chancy; in fiction, it is deterministic. No worries. The character in my example, of course, was meant to be someone who believed that the chance of heads was 50% in truth—not in fiction, however instrumentally useful such fiction might be.

There is no chance without chance. If our world is deterministic there are no chances in it, save chances of zero and one. Likewise if our world somehow contains deterministic enclaves, there are no chances in those enclaves. If a determinist says that a tossed coin is fair, and has an equal chance of falling heads or tails, he does not mean what I mean when he speaks of chance. Then what *does* he mean? This, I suppose, is the question Levi would like to see addressed. It is, of course, a more urgent question for determinists than it is for me.

That question has been sufficiently answered in the writings of Richard Jeffrey and Brian Skyrms on objectified and resilient credence.[4] Without commiting themselves one way or the other on the question of determinism, they have offered a kind of counterfeit chance to meet the needs of the determinist. It is a relative affair, and apt to go indeterminate, hence quite unlike genuine chance. But what better could a determinist expect?

According to my second formulation of the Principal Principle, we have the history-theory partition (for any given time); and the chance distribution (for any given time and world) comes from any reasonable initial credence function by conditionalizing on the true cell of this partition. That is, it is objectified in the sense of Jeffrey. Let us note three things about the history-theory partition.

(1) It seems to be a natural partition, not gerrymandered. It is what we get by dividing possibilities as finely as possible in certain straightforward respects.

[4] Richard C. Jeffrey, *The Logic of Decision* (New York: McGraw-Hill, 1965; second edition, Chicago: University of Chicago Press, 1983) Section 12.7; Brian Skyrms, "Resiliency, Propensities, and Causal Necessity," *Journal of Philosophy* 74 (1977): 704–13; Brian Skyrms, *Causal Necessity* (New Haven: Yale University Press, 1980).

(2) It is to some extent feasible to investigate (before the time in question) which cell of this partition is the true cell; but

(3) it is unfeasible (before the time in question, and without peculiarities of time whereby we could get news from the future) to investigate the truth of propositions that divide the cells.

Hence if we start with a reasonable initial credence function and do enough feasible investigation, we may expect our credences to converge to the chances; and no amount more feasible investigation (before the time) will undo that convergence. That is, after enough investigation, our credences become resilient in the sense of Skyrms. And our credences conditional on cells of the partition are resilient from the outset.

Conditions (1)–(3) characterize the history-theory partition; but not uniquely. Doubtless there are other, coarser partitions, that also satisfy the conditions. How feasible is feasible? Some investigations are more feasible than others, depending on the resources and techniques available, and there must be plenty of boundaries to be drawn between the feasible and the unfeasible before we get to the ultimate boundary whereby investigations that divide the history-theory cells are the most unfeasible of all. Any coarser partition, if it satisfies conditions (1)–(3) according to some appropriate standards of feasible investigation and of natural partitioning, gives us a kind of counterfeit chance suitable for use by determinists: namely, reasonable credence conditional on the true cell of that partition. Counterfeit chances will be relative to partitions; and relative, therefore, to standards of feasibility and naturalness; and therefore indeterminate unless the standards are somehow settled, or at least settled well enough that all remaining candidates for the partition will yield the same answers. Counterfeit chances are therefore not the sort of thing we would want to find in our fundamental physical theories, or even in our theories of radioactive decay and the like. But they will do to serve the conversational needs of determinist gamblers.

C. LAWS OF CHANCE

Despite the foundational problems of quantum mechanics, it remains a good guess that many processes are governed by probabilistic laws of nature. These laws of chance, like other laws of nature, have the form of universal generalizations. Just as some laws concern forces, which are magnitudes pertaining to particulars, so some laws concern single-case chances, which likewise are magnitudes pertaining to particulars.

For instance, a law of chance might say that for any tritium atom and any time when it exists, there is such-and-such chance of that atom decaying within one second after that time.[5] What makes it at least a regularity—a true generalization—is that for each tritium atom and time, the chance of decay is as the law says it is. What makes it a law, I suggest, is the same thing that gives some others regularities the status of laws: it fits into some integrated system of truths that combines simplicity with strength in the best way possible.[6]

This is a kind of regularity theory of lawhood; but it is a collective and selective regularity theory. Collective, since regularities earn their lawhood not by themselves, but by the joint efforts of a system in which they figure either as axioms or as theorems. Selective, because not just any regularity qualifies as a law. If it would complicate the otherwise best system to include it as an axiom, or to include premises that would imply it, and if it would not add sufficient strength to pay its way, then it is left as a merely accidental regularity.

Five remarks about the best-system theory of lawhood may be useful before we return to our topic of how this theory works in the presence of chance.

[5] Peter Railton employs laws of chance of just this sort to bring probabilistic explanation under the deductive-nomological model. The outcome itself cannot be deduced, of course; but the single-case chance of it can be. See Railton, "A Deductive- Nomological Model of Probabilistic Explanation," *Philosophy of Science* 45 (1978): 206–26; and the final section of my "Causal Explanation" in this volume.

[6] I advocate a best-system theory of lawhood in *Counterfactuals* (Oxford: Blackwell, 1973), pp. 73–75. Similar theories of lawhood were held by Mill and, briefly, by Ramsey. See John Stuart Mill, *A System of Logic* (London: Parker, 1843), Book III, Chapter IV, Section 1; and F. P. Ramsey, "Universals of Law and of Fact," in his *Foundations* (London: Routledge & Kegan Paul, 1978). For further discussion, see John Earman, "Laws of Nature: The Empiricist Challenge," in *D. M. Armstrong*, ed. by Radu J. Bogdan (Dordrecht: Reidel, 1984).

Mill's version is not quite the same as mine. He says that the question what are the laws of nature could be restated thus: "What are the fewest general propositions from which all the uniformities which exist in the universe might be deductively inferred?"; so it seems that the ideal system is supposed to be complete as regards uniformities, that it may contain only general propositions as axioms, and that its theorems do not qualify as laws.

It is not clear to me from his brief statement whether Ramsey's version was quite the same as mine. His summary statement (after changing his mind) that he had taken laws to be "consequences of those propositions we should take as axioms if we knew everything and organized it as simply as possible into a deductive system" (*Foundations*, p. 138) is puzzling. Besides Ramsey's needless mention of knowledge, his "it" with antecedent "everything" suggests that the ideal system is supposed to imply everything true. Unless Ramsey made a stupid mistake, which is impossible, that cannot have been his intent; it would make all regularities come out as laws.

(1) The standards of simplicity, of strength, and of balance between them are to be those that guide us in assessing the credibility of rival hypotheses as to what the laws are. In a way, that makes lawhood depend on us—a feature of the approach that I do not at all welcome! But at least it does not follow that lawhood depends on us in the most straightforward way: namely, that if our standards were suitably different, then the laws would be different. For we can take our actual standards as fixed, and apply them in asking what the laws would be in various counterfactual situations, including counterfactual situations in which people have different standards—or in which there are no people at all. Likewise, it fortunately does not follow that the laws are different at other times and places where there live people with other standards.

(2) On this approach, it is not to be said that certain generalizations are *lawlike* whether or not they are true, and the laws are exactly those of the lawlikes that are true. There will normally be three possibilities for any given generalization: that it be false, that it be true but accidental, and that it be true as a law. Whether it is true accidentally or as a law depends on what else is true along with it, thus on what integrated systems of truths are available for it to enter into. To illustrate the point: it may be true accidentally that every gold sphere is less than one mile in diameter; but if gold were unstable in such a way that there was no chance whatever that a large amount of gold could last long enough to be formed into a one-mile sphere, then this same generalization would be true as a law.

(3) I do not say that the competing integrated systems of truths are to consist entirely of regularities; however, only the regularities in the best system are to be laws. It is open that the best system might include truths about particular places or things, in which case there might be laws about these particulars. As an empirical matter, I do not suppose there are laws that essentially mention Smith's garden, the center of the earth or of the universe, or even the Big Bang. But such laws ought not to be excluded *a priori*.[7]

(4) It will trivialize our comparisons of simplicity if we allow our competing systems to be formulated with just any hoked-up primi-

[7] In defense of the possibility that there might be a special law about the fruit in Smith's garden, see Michael Tooley, "The Nature of Laws," *Canadian Journal of Philosophy* 7 (1977): 667–98, especially p. 687; and D. M. Armstrong, *What is a Law of Nature?* (Cambridge: Cambridge University Press, 1983), Sections 3.I, 3.II, and 6.VII. In "The Universality of Laws," *Philosophy of Science* 45 (1978): 173–81, John Earman observes that the best-system theory of lawhood avoids any *a priori* guarantee that the laws will satisfy strong requirements of universality.

tives. So I take it that this kind of regularity theory of lawhood requires some sort of inegalitarian theory of properties: simple systems are those that come out formally simple when formulated in terms of perfectly natural properties. Then, sad to say, it's useless (though true) to say that the natural properties are the ones that figure in laws.[8]

(5) If two or more systems are tied for best, then certainly any regularity that appears in all the tied systems should count as a law. But what of a regularity that appears in some but not all of the tied systems? We have three choices: it is not a law (take the intersection of the tied systems); it is a law (take the union); it is indeterminate whether it is law (apply a general treatment for failed presuppositions of uniqueness). If required to choose, I suppose I would favor the first choice; but it seems a reasonable hope that nature might be kind to us, and put some one system so far out front that the problem will not arise. Likewise, we may hope that some system will be so far out front that it will win no matter what the standards of simplicity, strength, and balance are, within reason. If so, it will also not matter if these standards themselves are unsettled. To simplify, let me ignore the possibility of ties, or of systems so close to tied that indeterminacy of the standards matters; if need be, the reader may restore the needed complications.

To return to laws of chance: if indeed there are chances, they can be part of the subject matter of a system of truths; then regularities about them can appear as axioms or theorems of the best system; then such regularities are laws. Other regularities about chances might fail to earn a place in the best system; those ones are accidental. All this is just as it would be for laws about other magnitudes. So far, so good.

But there is a problem nearby; not especially a problem about laws of chance, but about laws generally in a chancy world. We have said that a regularity is accidental if it cannot earn a place in the best system: if it is too weak to enter as an axiom, and also cannot be made to follow as a theorem unless by overloading the system with particular information. That is one way to be accidental; but it seems that a regularity might be accidental also for a different and simpler reason. It might hold merely by chance. It might be simple and powerful and well deserve a place in the ideal system and yet be no law. For it might have, or it might once have had, some chance of failing to hold; whereas it seems very clear, *contra* the best-system theory as so far stated, that no genuine law ever could have had any chance of not holding. A world of

[8] See my "New Work for a Theory of Universals," *Australasian Journal of Philosophy* 61 (1983): 343–77, especially pp. 366–68.

lawful chance might have both sorts of accidental regularitites, some disqualified by their inadequate contribution to simplicity and strength and others by their chanciness.

Suppose that radioactive decay is chancy in the way we mostly believe it to be. Then for each unstable nucleus there is an expected lifetime, given by the constant chance of decay for a nucleus of that species. It might happen—there is some chance of it, infinitesimal but not zero—that each nucleus lasted for precisely its expected lifetime, no more and no less. Suppose that were so. The regularity governing lifetimes might well qualify to join the best system, just as the corresponding regularity governing *expected* lifetimes does. Still, it is not a law. For if it were a law, it would be a law with some chance—in fact, an overwhelming chance—of being broken. That cannot be so.[9]

(Admittedly, we do speak of defeasible laws, laws with exceptions, and so forth. But these, I take it, are rough-and-ready approximations to the real laws. There real laws have no exceptions, and never had any chance of having any.)

Understand that I am not supposing that the constant chances of decay are *replaced* by a law of constant lifetimes. That is of course possible. What is not possible, unfortunately for the best-system theory, is for the constant chances to remain and to coexist with a law of constant lifetimes.

If the lifetimes chanced to be constant, and if the matter were well investigated, doubtless the investigators would come to believe in a law of constant lifetimes. But they would be mistaken, fooled by a deceptive coincidence. It is one thing for a regularity to be a law; another thing for it to be so regarded, however reasonably. Indeed, there are philosphers who seem oblivious to the distinction; but I think these philosophers misrepresent their own view. They are sceptics; they do not believe in laws of nature at all, they resort to regarded-as-law regularities as a substitute, and they call their substitute by the name of the real thing.

[9] At this point I am indebted to correspondence and discussion with Frank Jackson, arising out of his discussion of "Hume worlds" in "A Causal Theory of Counterfactuals," *Australasian Journal of Philosophy* 55 (1977): 3–21, especially pp. 5–6. A Hume world, as Jackson describes it, is "a possible world where every particular fact is as it is in our world, but there are no causes or effects at all. Every regular conjunction is an accidental one, not a causal one." I am not sure whether Jackson's Hume world is one with chances — lawless chances, of course — or without. In the former case, the bogus laws of the Hume world would be like our bogus law of constant lifetimes, but on a grander scale.

So the best-system theory of lawhood, as it stands, is in trouble. I propose this correction. Previously, we held a competition between all true systems. Instead, let us admit to the competition only those systems that are true not by chance; that is, those that not only are true, but also have never had any chance of being false. The field of eligible competitors is thus cut down. But then the competition works as before. The best system is the one that achieves as much simplicity as is possible without excessive loss of strength, and as much strength as is possible without excessive loss of simplicity. A law is a regularity that is included, as an axiom or as a theorem, in the best system.

Then a chance regularity, such as our regularity of constant lifetimes, cannot even be included in any of the competing systems. *A fortiori*, it cannot be included in the best of them. Then it cannot count as a law. It will be an accidental regularity, and for the right reason: because it had a chance of being false. Other regularities may still be accidental for our original reason. These would be regularities that never had any chance of being false, but that don't earn their way into the best system because they don't contribute enough to simplicity and strength. For instance suppose that (according to regularities that do earn a place in the best system) a certain quantity is strictly conserved, and suppose that the universe is finite in extent. Then we have a regularity to the effect that the total of this quantity, over the entire universe, always equals a certain fixed value. This regularity never had any chance of being false. But it is not likely to earn a place in the best system and qualify as a law.

In the paper, I made much use of the history-to-chance conditionals giving hypothetical information about the chance distribution that would follow a given (fully specified) initial segment of history. Indeed, my reformulation of the Principal Principle involves a "complete theory of chance" which is the conjunction of all such history-to-chance conditionals that hold at a given world, and which therefore fully specifies the way chances at any time depend on history up to that time.

It is to be hoped that the history-to-chance conditionals will follow, entirely or for the most part, from the laws of nature; and, in particular, from the laws of chance. We might indeed impose a requirement to that effect on our competing systems. I have chosen not to. While the thesis that chances might be entirely governed by law has some plausibility, I am not sure whether it deserves to be built into the analysis of lawhood. Perhaps rather it is an empirical thesis: a virtue that we may hope distinguishes our world from more chaotic worlds.

At any rate, we can be sure that the history-to-chance conditionals

will not conflict with the system of laws of chance. Not, at any rate, in what they say about the outcomes and chances that would follow any initial segment of history that ever had any chance of coming about. Let H be a proposition fully specifying such a segment. Let t be a time at which there was some chance that H would come about. Let L be the conjunction of the laws. There was no chance, at t, of L being false. Suppose for *reductio* first that we have a history-to-chance conditional "if H, then A" (where A might, for instance, specify chances at the end-time of the segment); and second that H and L jointly imply not-A, so that the conditional conflicts with the laws. The conditional had no chance at t of being false—this is an immediate consequence of the reformulated Principal Principle. Since we had some chance at t of H, we had some chance of H holding along with the conditional, hence some chance of H and A. And since there was no chance that L would be false, there was some chance that all of H, A, and L would hold together, so some chance at t of a contradiction. Which is impossible: there never can be any chance of a contradiction.

A more subtle sort of conflict also is ruled out. Let t, L, and H be as before. Suppose for *reductio* first that we have a history-to-chance conditional "if H, then there would be a certain positive chance of A"; and second that H and L jointly imply not-A. This is not the same supposition as before: after all, it would be no contradiction if something had a positive chance and still did not happen. But it is still a kind of conflict: the definiteness of the law disagrees with the chanciness of the conditional. To rule it out, recall that we had at t some chance of H, but no chance of the conditional being false; so at t there was a chance of H holding along with the conditional; so at t there was a chance that, later, there would be a chance of A following the history H; but chanciness does not increase with time (assuming, as always, the normal asymmetries); an earlier chance of a later chance of something implies an earlier chance of it; so already at t there was some chance of H and A holding together. Now we can go on as before: we have that at t there was no chance that L would be false, so some chance that all of H, A, and L would hold together, so some chance at t of a contradiction; which is impossible.

The best-system theory of lawhood in its original form served the cause of Humean supervenience. History, the pattern of particular fact throughout the universe, chooses the candidate systems, and the standards of selection do the rest. So no two worlds could differ in laws without differing also in their history. But our correction spoils that. The laws—laws of chance, and other laws besides—supervene now on the pattern of particular chances. If the chances in turn somehow supervene on history, then we have Humean supervenience of the laws as well; if not, not. The corrected theory of lawhood starts with the chances. It does nothing to explain them.

Once, *circa* 1975, I hoped to do better: to extend the best-system approach in such a way that it would provide for the Humean supervenience of chances and laws together, in one package deal. This was my plan. We hold a competition of deductive systems, as before; but we impose less stringent requirements of eligibility to enter the competition, and we change the terms on which candidate systems compete. We no longer require a candidate system to be entirely true, still less do we require that it never had any chance of being false. Instead, we only require that a candidate system be true in what it says about history; we leave it open, for now, whether it also is true in what it says about chances. We also impose a requirement of coherence: each candidate system must imply that the chances are such as to give that very system no chance at any time of being false. Once we have our competing systems, they vary in simplicity and in strength, as before. But also they vary in what I shall call *fit*: a system fits a world to the extent that the history of that world is a comparatively probable history according to that system. (No history will be very probable; in fact, any history for a world like ours will be very improbable according to any system that deserves in the end to be accepted as correct; but still, some are more probable than others.) If the histories permitted by a system formed a tree with finitely many branch points and finitely many alternatives at each point, and the system specified chances for each alternative at each branch point, then the fit between the system and a branch would be the product of these chances along that branch; and likewise, somehow, for the general, infinite case. (Never mind the details if, as I think, the plan won't work anyway.) The best system will be the winner, now, in a three-way balance between simplicity, strength, and fit. As before, the laws are the generalizations that appear as axioms or theorems in the best system; further, the true chances are the chances as they are according to the best system. So it turns out that the best system is true in its entirety— true in what it says about chances, as well as in what it says about history. So the laws of chance, as well as other laws, turn out to be true; and further, to have had no chance at any time of being false. We have our Humean supervenience of chances and of laws; because history selects the candidate systems, history determines how well each one fits, and our standards of selection do the rest. We will tend, *ceteris paribus*, to get the proper agreement

between frequencies and uniform chances, because that agreement is conducive to fit. But we leave it open that frequencies may chance to differ from the uniform chances, since *ceteris* may not be *paribus* and the chances are under pressure not only to fit the frequencies but also to fit into a simple and strong system. All this seems very nice.

But it doesn't work. Along with simpler analyses of chance in terms of actual frequency, it falls victim to the main argument in the last section of the paper. Present chances are determined by history up to now, together with history-to-chance conditionals. These conditionals are supposed to supervene, via the laws of chance of the best system, on a global pattern of particular fact. This global pattern includes future history. But there are various different futures which have some present chance of coming about, and which would make the best system different, and thus make the conditionals different, and thus make the present chances different. We have the actual present chance distribution over alternative futures, determined by the one future which will actually come about. Using it, we have the expected values of the present chances: the average of the present chances that would be made true by the various futures, weighted by the chances of those futures. But these presently expected values of present chances may differ from the actual present chances. A peculiar situation, to say the least.

And worse than peculiar. Enter the Principal Principle: it says first that if we knew the present chances, we should conform our credences about the future to them. But it says also that we should conform our credences to the expected values of the present chances.[10] If the two

[10] Let A be any proposition; let $P_1, P_2, \ldots$ be a partition of propositions to the effect that the present chance of A is $x_1, x_2, \ldots$, respectively; let these propositions have positive present chances of $y_1, y_2, \ldots$, respectively; let C be a reasonable initial credence function; let E be someone's present total evidence, which we may suppose to be presently admissible. Suppose that $C(-/E)$ assigns probability 1 to the propositions that the present chance of P_1 is y_1, the present chance of P_2 is $y_2, \ldots$. By additivity,

(1) $C(A/E) = C(A/P_1E)C(P_1/E) + C(A/P_2E)C(P_2/E) + \ldots$

By the Principal Principle,

(2) $C(P_1/E) = y_1,$

$C(P_2/E) = y_2,$

$\ldots$

and

differ, we cannot do both. So if the Principle is right (and if it is possible to conform our credences as we ought to), the two cannot differ. So a theory that says they can is wrong.

That was the strategy behind my argument in the paper. But I streamlined the argument by considering one credence in particular. Let T be a full specification of history up to the present and of present chances; and suppose for *reductio* that F is a nonactual future, with some positive present chance of coming about, that would give a different present distribution of chances. What is a reasonable credence for F conditionally on T? Zero, because F contradicts T. But not zero, by the Principal Principle, because it should equal the positive chance of F according to T. This completes the *reductio*.

This streamlining might hide the way the argument exploits a predicament that arises already when we consider chance alone. Even one who rejects the very idea of credence, and with it the Principal Principle, ought to be suspicious of a theory that permits discrepancies between the chances and their expected values.

If anyone wants to defend the best-system theory of laws and chances both (as opposed to the best-system theory of laws, given chances), I suppose the right move would be to cripple the Principal Principle by declaring that information about the chances at a time is *not*, in general, admissible at that time; and hence that hypothetical information about chances, which can join with admissible historical information to imply chances at a time, is likewise inadmissible. The reason would be that, under the proposed analysis of chances, information about present chances is a disguised form of inadmissible information about future history—to some extent, it reveals the outcomes of matters that are presently chancy. That crippling stops all versions of our *reductio* against positive present chances of futures that would

(3) $C(A/P_1E) = x_1$,

$C(A/P_2E) = x_2$,

...

(Since the $C(P_i/E)$'s are positive, the $C(A/P_iE)$'s are well defined.) So we have the prescription

(4) $C(A/E) = y_1x_1 + y_2x_2 + \ldots$

that the credence is to be equal to the expected value of chance.

yield different present chances.[11] I think the cost is excessive; in ordinary calculations with chances, it seems intuitively right to reply on this hypothetical information. So, much as I would like to use the best-system approach in defense of Humean supervenience, I cannot support this way out of our difficulty.

I stand by my view, in the paper, that if there is any hope for Humean supervenience of chances, it lies in a different direction: the history-to-chance conditionals must supervene trivially, by not being contingent at all. As noted, that would impose remarkably stringent standards on reasonable belief. To illustrate: on this hypothesis, enough purely historical information would suffice to tell a reasonable believer whether the half-life of radon is 3.825 days or 3.852. What is more: enough purely historical information *about any initial segment of the universe*, however short, would settle the half-life! (It might even be a segment before the time when radon first appeared.) For presumably the half-life of radon is settled by the laws of chance; any initial segment of history, aided by enough noncontingent history-to-chance conditionals, suffices to settle any feature of the world that never had a chance to be otherwise; and the laws are such a feature. But just how is the believer, however reasonable, supposed to figure out the half-life given his scrap of ancient history? We can hope, I suppose, that some appropriate symmetries in the space of possibilities would do the trick. But it seems hard to connect these hoped-for symmetries with anything we now know about the workings of radioactive decay!

D. RESTRICTED DOMAINS

In reformulating the Principal Principle, I took care not to presuppose that the domain of a chance distribution would include all propositions. Elsewhere I was less cautious. I am grateful to Zeno Swijtink for

[11] As to the version in the paper: declaring hypothetical information about chances inadmissible blocks my reformulation of the Principal Principle, and it was this reformulation that I used in the *reductio*.

As to the version in the previous footnote: if information about present chances is inadmissible, then it becomes very questionable whether the total evidence E can indeed be admissible, given that $C(-/E)$ assigns probability 1 to propositions about present chance.

As to the streamlined version in this postscript: T includes information about present chances, and its partial inadmissibility would block the use of the Principal Principle to prescribe positive credence for F conditionally on T.

pointing out (personal communication, 1984) that if I am to be uniformly noncommital on this point, two passages in my final section need correction.

I say that if C_1 and C_2 are any two reasonable initial credence functions, and Y is any member of the history-theory partition for any time, then $C_1(-/Y)$ and $C_2(-/Y)$ are "exactly the same." Not so. The most I can say is that they agree exactly in the values they assign to the propositions in a certain (presumably large) set; namely, the domain of the chance distribution implied by Y. My point stands: I have a consequence of the Principal Principle that is entirely about credence, and that limits the ways in which reasonable initial credence functions can differ.

Later I say that these differences are—implausibly—even more limited on the hypothesis that the complete theory of chance is the same for all worlds. The same correction is required, this time with complete histories in place of history-theory conjunctions. Again my point stands. The limitation of difference is less than I said, but still implausibly stringent. Unless, of course, there are very few propositions which fall in the domains of chance distributions; but that hypothesis also is very implausible, and so would not save the day for a noncontingent theory of chance and for Humean supervenience.

My reason for caution was not that I had in mind some interesting class of special propositions—as it might be, about free choices—that would somehow fail to have well-defined chances. Rather, I thought it might lead to mathematical difficulties to assume that a probability measure is defined on all propositions without exception. In the usual setting for probability theory—values in the standard reals, sigma-additivity—that assumption is indeed unsafe: by no means just any measure on a restricted domain of subsets of a given set can be extended to a measure on all the subsets. I did not know whether there would be any parallel difficulty in the nonstandard setting; it probably depends on what sort of infinite additivity we wish to assume, just as the difficulty in the standard setting arises only when we require more than finite additivity.

Plainly this reason for caution is no reason at all to think that the domains of chance distributions will be notably sparser than the domains of idealized credence functions.

· TWENTY ·

Probabilities of Conditionals and Conditional Probabilities

The truthful speaker wants not to assert falsehoods, wherefore he is willing to assert only what he takes to be very probably true. He deems it permissible to assert that A only if $P(A)$ is sufficiently close to 1, where P is the probability function that represents his system of degrees of belief at the time. Assertability goes by subjective probability.

At least, it does in most cases. But Ernest Adams has pointed out an apparent exception.[1] In the case of ordinary indicative conditionals, it seems that assertability goes instead by the conditional subjective probability of the consequent, given the antecedent. We define the conditional probability function $P(-/-)$ by a quotient of absolute probabilities, as usual:

(1) $P(C/A) = \mathrm{df}\ P(CA)/P(A)$, if $P(A)$ is positive.

(If the denominator $P(A)$ is zero, we let $P(C/A)$ remain undefined.) The truthful speaker evidently deems it permissible to assert the indicative conditional that if A, then C (for short, $A \rightarrow C$) only if $P(C/A)$ is

[1] Ernest Adams, "The Logic of Conditionals", *Inquiry* 8 (1965), 166–197; and "Probability and the Logic of Conditionals", *Aspects of Inductive Logic*, ed. by Jaakko Hintikka and Patrick Suppes, Dordrecht, 1966. I shall not here consider Adams's subsequent work, which differs at least in emphasis.