PROPENSITIES AND PRAGMATISM[*]

Abstract: A pragmatist conception of propensity is outlined, and defended against common objections to the propensity interpretation of probability, prominently Humphreys' paradox. The paradox is reviewed, and one of its key assumptions is identified. The *identity thesis* states that propensities are probabilities (under a suitable interpretation of Kolmogorov's axioms). It is argued that the *identity thesis* is involved in many empiricist versions of the propensity interpretation deriving from Popper's original and influential proposal, and is one of the main reasons why such interpretations are untenable. As an alternative, a return to Charles Peirce's insights on probabilistic dispositions is recommended, and a reconstructed version of his pragmatist conception is offered, which rejects the *identity thesis*. – Correspondence to: msuarez@filos.ucm.es

The propensity interpretation of probability was introduced by Karl Popper in a series of epoch-making papers in the late 1950s. However, the more general thought that dispositional properties are essentially connected to chance had been voiced earlier: Charles Peirce is often credited with the introduction of the main insight as far back as

[*] Forthcoming in *The Journal of Philosophy*, February 2013 (submitted 6 June, accepted 23 November, 2011). For comments I would like to thank the audience at the III Graduate Conference in the Philosophy of Probability, London School of Economics (June 2010), and Donald Gillies and Alan Hajek. The paper was written during research stays at Harvard University and Oxford in 2010-11. At Harvard I thank the Philosophy Department for sponsorship; at Oxford, I thank the Bodleain Library. Financial support is acknowledged from the Spanish Government research grants FFI2008-06418-C03-01, FFI2011-29834-C03-01 and PR2010-0066.

1892. One of my claims in this paper is that there are substantial differences in the spirit, as well as the details, of Popper's and Peirce's accounts – and that these differences matter in the evaluation of some contemporary claims and arguments in the philosophy of probability. Their accounts may be referred to as the 'empiricist' and 'pragmatist' conceptions of propensity respectively, [1] and in this paper I argue for a reconstructed version of the pragmatist conception.

When it comes to probability, the empiricist tradition has traditionally favored the frequency interpretation as formulated by Reichenbach and Von Mises. [2] On this view probability is in some sense an extension of the concept of statistical association or correlation among observable quantities. More precisely, probability is identified with the ratio or frequency of favourable to total outcomes in a sequence of results of a repeated chance experiment. The sequences are defined with respect to some particular reference class – or "collective" in Von Mises terminology. The view is supposedly in line with a Humean understanding of laws as regularities, and more particularly of probabilistic laws as statistical regularities.

By contrast, Popper's introduction of the propensity interpretation may *prima facie* appear to be a nonempiricist, or even metaphysical, retort to the frequency interpretation. Certainly, Popper was not an empiricist in any traditional sense. He was

---

[1] There are empiricist and pragmatist elements in both Peirce's and Popper's philosophies of probability. It is moreover arguable that their overall philosophies of science – their scientific methodologies in particular - are very close in many respects. But the labels capture well the relative differences between their respective views on probability, and fittingly convey the fundamental difference in their spirit.

[2] Hans Reichenbach, *The Theory of Probability: An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability* (Los Angeles: University of California Press, 1935/49). Richard Von Mises, *Probability, Statistics and Truth* (New York: Dover, 1957).

neither an inductivist nor a verificationist. He did not think the content of scientific knowledge extends only as far as our knowledge of the empirical or phenomenological realm. He liked instead to describe himself as a critical rationalist. He was convinced that metaphysical speculation was an important part of the creative process that generates scientific conjectures. And he was not opposed to the postulation of ontological entities as part of scientific explanations and theories.

The main defect that Popper identified in the frequency interpretation is related to its insurmountable difficulties in accounting for single-case probabilities, which Popper thought were necessary for quantum mechanics. Suppose we carry out a chance experiment on a system only once (for example, tossing a coin). An outcome event is produced (for example, heads), but there is then no well-defined reference class or collective with respect to which we may define a sequence; and there is thus no frequency within a sequence that we may identify the probability of the event with. So, on a frequency interpretation, the probability of a single event is meaningless or undefined.

The propensity interpretation solves this problem simply by postulating the existence of a propensity (e.g. a ½ propensity to heads) that is well defined and meaningful independently of any collective. The move to postulate "propensities" over and above any features of any sequences is strikingly in violation of the spirit of Humeanism, since it blocks the possibility of a reduction of probabilities to frequencies. Not surprisingly Humeans roundly reject propensities.

The conflict between propensities and Humeanism is sometimes mistakenly thought to extend to all forms of empiricism. Contemporary versions of the propensity interpretation make it clear that this extension is illegitimate. In particular Donald Gillies has developed his propensity interpretation in the context of a broad empiricist methodology. [3] And indeed in a wide-enough sense of the word Popper too was an empiricist. He defended the primacy of the scientific method and its rationality. And he took the empirical sciences to be primarily characterised by their close connection to empirical evidence. The testing of hypotheses in particular was a key and essential part of the scientific method. His critical engagement with Vienna Circle logical empiricism is very much at the heart of his philosophical development, as he continued to quarrel and debate with their ideas and those of their disciples throughout his life. [4] Given this critical engagement it cannot be surprising that some elements of their empiricism crept into aspects of Popper's philosophy, including his philosophy of probability. One of these elements in particular, I argue, makes it legitimate to refer to Popper's propensity interpretation as "empiricist".

By contrast, Peirce's views on chance were not prompted by any physical theory, or any particular empirical findings. He was instead led to dispositional notions by his acquaintance and practical engagement with techniques of statistical inference in geology and astronomy. [5] His approach was pragmatist in a broad sense that accords to

---

[3] Donald Gillies, *Philosophical Theories of Probability* (London: Routledge, 2000b).
[4] For an account, see Victor Kraft "Popper and the Vienna Circle", in Philip A. Schilpp (ed.), *The Philosophy of Karl Popper* (La Salle Illinois: Open Court, 1974), pp. 185-204, which suitably emphasises Popper's empiricism.
[5] Peirce's work as an employee of the United States Coast and Geodetic Survey is well documented – see Victor F. Lenzen, "Charles S. Peirce as Astronomer", in Edward C. Moore and Richard Robin (eds.), *Studies in the Philosophy of Charles Sanders Peirce* (Armherst, Massachusetts: University of Massachusetts Press, 1964), pp. 33-50, and Ian

his own philosophy: he did not reject ontological postulates as long as they exhibited explanatory power, and he thought inference to the best explanation may well ground theoretical dispositional properties. He moreover worked hard to understand the practical consequences or manifestations of probabilistic dispositions, and made an effort to figure out what difference chance ascriptions actually make in practice.

Section 1 reviews Popper's propensity interpretation of probability and identifies a residual empiricist commitment – which I call the *identity thesis*. Sections 2 and 3 review Peirce's notion of probabilistic dispositional property in the context of his pragmatist philosophy. Section 4 argues for an updated version of the 'pragmatist' conception that rejects the identity thesis. Then in section 5 I turn to the contemporary debate by considering in detail a fundamental objection to the propensity interpretation of probability due to Paul Humphreys. Section 6 discusses some of the standard moves in response to Humphreys' 'paradox'. I argue that the identity thesis is the key to all of them, since it is a covert assumption in the formulation of the paradox itself. Section 7 develops the pragmatist conception in response to Humphreys' paradox. Section 8 concludes by defending the pragmatist conception against the most sophisticated empiricist account nowadays available.

1. Popper's Empiricist Propensity Interpretation

For our present purposes four elements in Popper's propensity interpretation stand out. First, in spite of Popper's occasional use of the word "theory", his proposal is to provide

---

Hacking, *The Taming of Chance* (Cambridge: Cambridge University Press, 1990), chapter 23.

an *interpretation* of the probability calculus. [6] Early on in his career, Popper considered alternative axiomatic systems, and even developed one of his own. [7] But by the time he developed his propensity views in the 1950's and 1960's, Kolmogorov's axioms were established. Thus Popper's application of propensity to probability was circumscribed by these. Second, Popper's interpretation is what I will call a *relational* one: it postulates propensities as properties of entire experimental set ups, not individual systems. On such views it makes no sense to suppose that an object in a single-object universe, with no further experimental apparatus, possesses any propensities. Popper writes: "A statement about propensity may be compared with a statement about the strength of an electric field […] and just as we can consider the field as physically real, so we can consider the propensities as physically real. They are relational properties of the experimental set-up." [8]

A third element is indeed Popper's empirical realism concerning propensities. Popper emphasised how propensities are part of empirical reality just like forces and masses are part of the empirical reality described by Newtonian mechanics. Propensities are hence neither fictional nor hypothetical: Their existence is certainly testable, and Popper thought that it had in fact been tested. [9]

---

[6] 'Interpretation' is taken in this paper in its most common usage in analytical philosophy – as a 'model', and therefore an account of the nature of something. To interpret *x* is to provide a model *m* for *x* – where *m* tells us what *x* really is.

[7] Karl Popper, "A Set of Independent Axioms for Probability", *Mind*, XLVII, pp. 275-277.

[8] Karl Popper, "The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory" in Stephan Körner (ed.), *Observation and Interpretation in the Philosophy of Physics* (New York: Dover, 1957), pp. 65-70.

[9] By interference experiments in quantum mechanics – see Karl Popper, "The Propensity Interpretation of Probability", *British Journal for the Philosophy of Science*, 37 (1959), p. 28; and Karl Popper, *Quantum Theory and Schism in Physics* (London: Hutchinson, 1982), pp. 83-84.

The fourth and final element is the adoption by Popper of a long-run version of the propensity interpretation, as opposed to a single-case version. [10] We may distinguish these views as follows. All propensity theories postulate dispositional properties that yield the appropriate empirical frequencies observed in experimental trials conducted on suitable chance set-ups. A propensity then, like any other dispositional property, manifests itself in the display of another distinct property, and this is true both in the long-run and single-case views. However, the nature of the manifestation property differs fundamentally on both views. A long-run view assumes that a propensity is a property of something like a repeatable sequence of experimental trials. So it may only manifest itself as a frequency in the sequence of outcomes of a series of experimental trials *in the long run*. A single-case view, on the other hand, assumes that the propensity is a property of a single trial and thus may manifest itself fully in that very trial. The former manifestation property belongs to the long-run sequence, while the latter is a property of the single trial.

The long-run view may be characterised as follows: A propensity to a particular outcome is a probability, corresponding to a distribution over the possible outcomes of an experimental chance set up, and giving rise to a frequency when the experiment is often repeated. [11] To give a routine example: Suppose that I toss a coin, which I know to

[10] There is some debate over the nature of Popper's interpretation at this point. His insistence on applying propensities to the single case often makes it sound as if he adopts a single-case interpretation. But in developing the details of the account he links propensities to repeatable conditions – something seemingly incompatible with a single-case theory. Later on in life he seems to have recognised this, and some authors even distinguish an "earlier" and a "later" Popper – e.g. Donald Gillies, op. cit. (2000b), pp. 126-29.

[11] I am at this point reformulating Donald Gillies' characterisations slightly in order to avoid their ambiguity over the distinction between what I call 'objectual' and 'relational' dispositions (cf. the definitions in Gillies, op. cit. (2000b), p. 126 and p. 131, for illustrations of the ambiguity).

be fair. Its probability of landing heads is ½. On Popper's propensity interpretation, this probability is a dispositional property of the generating conditions, or experimental setup, and it gives rise to a 50-50 frequency when the experiment is repeated infinitely. In the long-run version of this interpretation, contrary to what happens in the single-case version, the frequency *is* the property that displays the propensity; and there is no need to invoke any further entities or properties of the single trial.

The empiricist component in Popper's propensity interpretation is then best summarised in what we may call the Identity Thesis: the view that propensities *are* probabilities (under the appropriate interpretation of the probability calculus). This is an empiricist thesis in the sense that it links by definition an abstract explanatory concept ("propensity") to a more concrete concept in empirical science ("probability"). This empiricist commitment is expressed in two of the aspects of Popper's proposal discussed above. First, there is the insistence on propensities as interpretations of probabilities. That is, Popper is not postulating a set of distinct new entities *over and above* probabilities; he is rather providing a theory of probabilities *as* propensities. Second, there is the emphasis on considering these probabilities, so interpreted, as part of empirical reality, and thus subject to the same standards of confirmation as forces, masses or the like. Propensity ascriptions are, for Popper, testable in just the same way any other empirical conjectures are. Every empirical refutation of a statistical law vindicates their empirical character. There is therefore a legitimate sense in which Popper's propensity interpretation may be said to be 'empiricist'.

2. Peirce's Chances

Charles Peirce is among those credited with first introducing the notion of absolute chance, or irreducible metaphysical indeterminism. His views on dispositional probabilities track those on absolute chance, and may even have preceded them. [12] What has not been sufficiently appreciated is how Peirce's philosophy of probability is in line with his general pragmatism – or 'pragmaticism' as he used to call it. It is standard nowadays to refer to any theory of probabilistic dispositions as a 'propensity' theory. Peirce never used the term 'propensity', and it would be anachronistic to ascribe it to him; but it may help us to compare Peirce's view to other contemporary 'propensity' theories. Indeed Peirce's propensities possess four distinguishing features; they are 'objectual', 'causal', 'hypothetical', and 'long run'.

Let me begin by quoting two well-known paragraphs in Peirce's writings:

> "[…] The die has a certain 'would-be'; and to say that a die has a 'would-be' is to say that it has a property, quite analogous to any habit that a man might have. […] And just as it would be necessary, in order to define a man's habit, to describe how it would lead him to behave and upon what sort of occasion – albeit this statement would by no means imply that the habit consists in that action – so to define the die's 'would-be' it is necessary to say how it would lead the die to behave on an occasion that would bring out the full consequence of the 'would-be'; and this statement will not of itself imply that the 'would-be' of the die consists in such behavior."

[12] As well as his logic of statistical inference – see Ian Hacking (op. cit., pp. 207-210).

> "Now in order that the full effect of the die's 'would-be' may find expression, it is necessary that the die should undergo an endless series of throws from the dice box, the result of no throw having the slightest influence upon the result of any other throw, or, as we express it, the throws must be independent each of every other." [13]

Together these two paragraphs express at least three of the four main features of Peirce's 'propensities'. First, notice that Peirce ascribes the propensity (the 'would-be' property) to the die itself, taken in isolation. In other words, Peirce holds an 'objectual' view that ascribes propensities to the chancy objects themselves. While this might at first sight appear trivial, it is as a matter of fact very unusual nowadays, for at least two reasons. The first reason is that contemporary theories typically ascribe propensities to events and not objects. (In the next section I argue that contemporary theories tend to identify propensities and probabilities, which means that to be coherent they must ascribe propensities to the elements of a sigma field, that is, events or propositions). The second reason is that the usual assumption nowadays is to ascribe dispositional properties to the entire experimental set up, which includes the chancy object but also a number of further entities and their properties too including, possibly, on some accounts, the state of the whole universe at a given time. However, these are arguably grand metaphysical departures from the common-sense or pre-philosophical view. For what is usual in common parlance in everyday life is to ascribe chance to the chancy object itself just as Peirce does. (A few examples: "this coin's chance of landing heads", "Arsenal's chance to win the Champions' League this year", "given this material's chance to radioactively decay in the next hour"). Peirce's 'objectualism' is thus grounded upon common dispositional parlance; it is the subsequent propensity theories

[13] Charles S. Peirce, "Note on the Doctrine of Chances" (1910), in *Philosophical Writings of Peirce* (New York: Dover, 1955), p. 169.

(including Popper's) which seem intriguingly at odds with the ordinary linguistic practice.

Second, Peirce's "propensities" are causal in the sense that he does not refrain from ascribing causal powers to chances. Notice in this regard the following locutions as applied to propensities in the quoted excerpts: they 'lead' [to a certain behaviour], and they have 'full consequences' that are 'brought out' on particular occasions. Peirce was clear in his writings about the nature of causation in general, which he thought of in terms of Aristotle's efficient causation, and distinguished carefully from the other three Aristotelian notions of causation, from Kant's notion of causation as the instantaneous determination of states, and from Hume's notion of causation as constant conjunction. [14] He then went on to explicitly link chances' powers to efficient causation. [15]

The 'hypothetical' character of Peirce's propensities is not explicit but it may be elicited from these paragraphs. Consider in particular Peirce's close analogy, in the quote given, between objective chances and human habits. Unlike other prima facie categorical properties of human beings, such as height or weight, habits are not susceptible to observation, measurement or perception through the senses. Our knowledge of such properties is rather the result of inference from testimony and observed behaviour. They tend to appear not in description of behaviour, but in its explanation and justification, as well as other self-vindicating practices. And just as habits are invoked in the explanation and justification of behaviour, chances are invoked in the explanation and

[14] Charles S. Peirce "Causation and Force" (1898), in *Collected Papers of Charles Sanders Peirce*, vol. VI: *Scientific Metaphysics*, edited by Charles Hartshorne and Paul Weiss (Cambridge: Cambridge University Press, 1935), pp. 46 – 87.
[15] Charles S. Peirce "Reply to the Necessitarians" (1893), in op. cit. (1935), appendix A, pp. 403ff.

justification of phenomena. In drawing a close analogy between habits and propensities, Peirce implicitly suggests that 'propensities' may be regarded as theoretical explanatory properties arrived at by the 'method of hypothesis'.

The second quoted paragraph also expresses Peirce's commitment to what is nowadays known as a 'long-run' propensity view. While avoiding empiricist strictures, Peirce's pragmatism compels him nonetheless to consider what difference propensity ascriptions would make in practice. And he finds that the full consequences of propensities can only be revealed in a long-run (virtual and infinite) sequence of experimental trials of the same kind. In more contemporary terms, propensities manifest themselves in the appropriate repeatable experiments as virtual or hypothetical limiting frequencies.

3. Maxims of Pragmatism

In this paper I argue for a reconstructed version of Peirce's view on probabilistic dispositions. Although my account differs from Peirce's original view in a number of respects, it is nonetheless squarely in Peirce's pragmatist tradition. In fact I claim that the most significant point of difference with Peirce actually serves to further advance the aims of pragmatism. Moreover, the contrast with Popper's 'empiricist' view becomes clearest against the background of a general pragmatist philosophy of science. So in this section I expound on this philosophical framework, by characterising it as clearly as possible, however tentatively, in terms of five maxims. We may refer to them collectively as the maxims of pragmatism, although strictly speaking only the first one was ever known under such a name. The remaining four, expressed here in the form of negative commands, are meant as succinct expressions of recurrent themes throughout

the history of pragmatism that have gained relevance in the light of subsequent philosophical developments. As foil and way of contrast I also deploy different theses from various empiricist traditions – in a brief summary form – indicating its provenance in each case. Such presentation is anachronistic, but it has the analytical virtue to bring into relief the key distinguishing features of these philosophical approaches to propensities. It does not aim at historical or exegetical accuracy, but rather to abstract away from unnecessary detail.

Maxim one (the pragmatist maxim): Do refer your concept of any object to what you conceive to be that object's effects, or practical consequences.

Peirce stated the pragmatist maxim in several different forms and with different emphases throughout his life. [16] In particular he seems to have slightly modified the relative importance he attached to 'practical consequences' as opposed to merely effects of the object. The one aspect that seems to have remained constant throughout was the emphasis on 'conceivability': The concept of the object is only exhausted by its full set of conceivable effects. In other words, the pragmatist maxim applies to all objects, whether actual, possible, or merely imaginary. And it defines any such object in terms

[16] Charles S. Peirce, "How to make our ideas clear" (1878), in *Writings of C. S. Peirce: A Chronological Edition*, edited by M. Fisch (Bloomington: University of Indiana Press, 1986), vol. 3, p. 266: "Consider what effects, which might conceivably have practical bearings, we conceive the object of our conception to have. Then our conception of these effects is the whole of our conception of the object". It is remarkable that the example that Peirce invokes straightaway is precisely the concept of a dispositional property, namely 'hardness'. The passage continues as follows: "Let us illustrate this rule by some examples; and to begin with the simplest one possible, let us ask what we mean by calling a thing *hard*. Evidently that it will not be scratched by many other substances. The whole conception of this quality, as of every other, lies in its conceived effects. There is absolutely no difference between a hard thing and a soft thing so long as they are not brought to the test".

of all its effects, whether actual, possible or merely imaginary – thus including (but not restricted to) those effects of the object in the actual world.

The remaining four maxims are compatible with, while not logically following from, maxim one. For our purposes they are best expressed negatively, in order to emphasise the contrast with empiricism, as follows.

Maxim two: Do not identify hypothetical entities and their properties with their empirical manifestations or displays.

The second maxim signals pragmatism's opposition to two distinct forms of empiricism, namely, verificationism and operationalism, at least in their crudest versions. The pragmatist rejects the need for a reduction of our abstract concepts to any phenomenological basis, such as sense-data, or any set of experiences. He or she instead assumes that the entities postulated in scientific theories stand on their own feet and require no grounding in empirical data in order to be meaningful. The truth conditions of statements regarding such entities and properties are not reducible to those of statements regarding their empirical verification or manifestations. Note that maxim two is compatible with maxim one. The latter is far more general, containing no reference to 'hypothetical' entities, or 'empirical' effects. For example, some of the conceivable effects of some of the unobservable entities postulated in scientific theories are themselves unobservable.

Maxim three: Do not seek mere analyses or philosophical interpretations of scientific concepts, but actively engage with the sciences, by refining, changing or adding to their concepts.

Pragmatists regard mere conceptual analysis as an unduly limiting task for philosophy, which is moreover underwritten by a mistaken general conception of both the nature of knowledge and its acquisition – the so-called spectator conceptions of knowledge. [17] On such conceptions philosophy is a distinct meta-discipline that can at best aim to clarify some of the concepts employed in the sciences – but it neither disputes them nor engages in any critical way with them. The pragmatist tradition, by contrast, has tended to see philosophy as continuous with science, engaging in a critical manner with both its methods and concepts, and bringing its own specific techniques to bear on scientific problems directly. And although philosophy may be able to provide a wider, bird's eye point of view on the state of the diverse sciences, it is not a fundamentally distinct form of inquiry.

This maxim opposes pragmatism to two distinct traditions. On the one hand it opposes it to the type of ordinary language philosophy that accepts conceptual analysis as the only legitimate form of philosophical reflection upon science. On the other hand it opposes pragmatism to views of the method and aim of philosophy at odds with those of science. The former distinguishes pragmatism from the tradition of conceptual analysis, while the latter distinguishes it from an array of methodological empiricisms, including Popper's own. Popper was the most reluctant among classical twentieth-century

[17] John Dewey, *The Quest for Certainty*, in J. Boydston (ed.), *John Dewey: The Later Works, 1925-1953*, vol. 4 (Carbondale and Edwardsville: Southern Illinois University Press, 1929), p. 19.

methodologists to subject his own philosophical proposals to the very methods that they postulated for science. Indeed the difficulties in applying Popper's proposed scientific method ('falsificationism') to itself are notorious, and the debate over the scientific credentials of Popper's philosophy of critical rationalism rages on to this day. [18]

Maxim four: Do not attempt to reduce causal efficacy and causation to anything empirically accessible, such as frequency or correlation.

Another form of empiricism opposed by pragmatism is Humeanism. Hume proposed to banish all metaphysical necessity from our knowledge of the world, and aimed to reduce all modalities to correlations or regularities among actual occurrences. Causation was one of his main targets, and many exegetical analyses of Hume's writings on causality agree that Hume attempted to reduce causation to contingent features of the actual world such as actual constant conjunction or, alternatively, the actual psychological habits by which humans come to expect those regularities to persist in time. [19] The pragmatist tradition on the whole has eschewed Humeanism along with any other attempt to reduce our causal talk to anything else. Peirce, in particular, freely employed the language of causes, tendencies, and powers, while not attempting to provide any theory that would reduce such concepts to more basic, elementary, or empirically accessible ones. Unlike some of their contemporaries in logical positivism,

[18] See e.g. Robert Nola and Howard Sankey, *Theories of Scientific Method* (London: Acumen, 2007).

[19] Thus following, roughly, Hume's two definitions of cause in the *Treatise* (David Hume, *A Treatise on Human Nature* (Oxford: Oxford University Press, 1739 / 1978), p. 77, and p. 170; see Robert Fogelin, *Hume's Skepticism ain the Treatise of Human Nature* (London: Routledge, 1985), chapter IV for discussion). Galen Strawson refers to the first form of reduction as "the standard view" in the exegesis of Hume's views on causation (in *The Secret Connection: Causation, Realism and David Hume* (Oxford: Oxford University Press, 1989), p. vii).

logical empiricism or logical atomism, the pragmatists were not at any stage inclined towards a foundationalist epistemology seeking to ground our knowledge of the external world upon the 'secure' foundations of the senses. The pragmatist conception of experience is notably wider than the empiricist one, and it crucially does not assume that a definite gradient of epistemic 'security' or 'certainty' attaches to claims exclusively reporting the content of our perceptions. [20] As with maxim two, notice that the denial of maxim four fails to follow from maxim one, which is more generally applied to all our concepts, and does not mark out an 'empirically accessible' basis for any reduction. For example, the practical consequences, or effects, of causes may well include further causes.

Maxim five: Do not systematically reject as unacceptable those scientific theories that postulate hypothetical or fictional entities.

Peirce was instrumental in establishing the common view nowadays that the sciences progress by means of ampliative forms of reasoning going well beyond inductive or deductive inference. Abduction, or the method of hypothesis, was outstanding among them; and under the name 'inference to the best explanation' it is at the centre of present-day debates over scientific realism. It is unclear whether Peirce demanded that theories arrived at by means of these ampliative leaps of reasoning should be true, at

[20] This is at any rate the spirit that notably underwrites Quine's critique of Carnap's notion of analyticity ("Two Dogmas of Empiricism", *Philosophical Review*, 60 (1951), pp. 20-43); Sellar's critique of the myth of the given ("Empiricism and the Philosophy of Mind" in Herbert Feigls and Michael Scriven (eds.), *The Foundations of Science and the Concepts of Psychoanalysis*, Minnesota Studies in the Philosophy of Science, vol. I, Minneapolis: University of Minnesota Press); and Neurath's critique of foundationalist epistemology (Otto Neurath, "Protocol Statements" in Robert Cohen and Marie Neurath (eds.), *Otto Neurath: Philosophical Papers* (Dordrecht: Reidel, 1932 / 1983).

least in the long term of ideal inquiry. He is certainly often interpreted that way. [21] At any rate, regardless of the merits of this interpretation, Peirce did not think that only a true theory is acceptable *now* or, indeed, at any (finite) stage in the development of science. On the contrary, for Peirce scientific theories postulating hypothetical or fictional entities are acceptable as long as they are explanatory – and explanatory power is at best a fallible guide to truth. Such theories may only, if anything, be rejected in the long term of an ideally conducted inquiry. Hence for a pragmatist the introduction of a fictional or hypothetical entity as part of a scientific theory or model may be justified on ampliative explanatory grounds. By contrast the empiricist only has to hand the tools of deductive or inductive reasoning in order to justify such postulates. Strict empiricists would reject fictional or hypothetical entities since they have not been derived from our secure perceptual knowledge by inductive means. And in Popperian methodology, the postulate of a new type of entity must at least in principle be empirically testable, for the scientific theory that introduces such an entity must be falsifiable. Neither is able to otherwise countenance the postulation of explanatory entities arrived at by the method of hypothesis.

4. Propensity: A Pragmatist Conception

In this section I explicitly compare Peirce's and Popper's views with respect to the five maxims of pragmatism described in the last section. The conclusion will be that the four features discussed of Peirce's dispositional probabilities are in agreement with the pragmatist maxims, while at least one important feature of Popper's propensities is not.

---

[21] Hilary Putnam, in *Reason, Truth and History* (Harvard: Harvard University Press, 1981), and Christopher Hookway, in *Peirce* (London: Routledge, 1985), interpret Peirce as entertaining a regulative-ideal conception of inquiry and, correspondingly, of scientific truth.

The key idea is related to the distinction between long-run and single-case versions of the propensity theory. As was noticed, the 'long-run' version requires a bipartite distinction between frequencies, on the one hand, and propensities on the other. The identity thesis then interprets probabilities as propensities, so the propensity theory is then regarded as an elucidation of what the concept of (objective) probability amounts to. The idea is that objective probabilities are somehow connected with dispositional properties – but how exactly? The contemporary theory of dispositional properties distinguishes appropriately between the possession and manifestation conditions of a disposition. We may think of dispositions as the underlying properties defined by the possession conditions; there must in addition be some manifestation property defined by the manifestation conditions. [22] It is natural then to think of propensities as the underlying dispositional properties of things with frequencies as their empirical manifestations. The coin's ½ chance is manifested in the long-run infinite virtual sequence of experimental trials of tossing coins. It is only in this long-run infinite virtual sequence that we may expect the relative ratio of heads and tails to be one. It follows that – as Peirce thought – the difference that a propensity ascription can make in practice can only be ascertained in a hypothetical long run of the experiment. (And, in fact, in the infinite long run since, as is well known, any finite relative frequency is compatible with any value in the ascription of a propensity). This must mean that for Peirce the practical consequences of a propensity ascription are as hypothetical as the propensity itself.

[22] Stephen Mumford, *Dispositions* (Oxford: Oxford University Press, 1998). In the typical example "fragile" is the dispositional property, and "breakage" is its manifestation.

It follows that Peirce did not hold the identity thesis. His aim could not have been to explicate, in Carnap's sense of the term, the concept of probability, as formally expressed by Kolmogorov. This is a project that he would not have even comprehended. He neither had to hand a particular formalization of the concept, nor did he share in the project of 'explicating' scientific concepts. We cannot seriously ascribe to him the empiricist commitment to replace a theoretical notion with an empirically accessible one. His diatribes against Hume on this particular matter are unambiguous. [23] And the fact that, as we just saw, he took both propensities and their manifestations to be hypothetical is at odds with any empiricist account of these notions. What Peirce was rather aiming for was a development of a causal understanding of chance and its practical consequences for doing science. He thought the concept of chance, so understood, was an integral part of the activity of scientists everywhere. He took it to be exemplified in most experimental physics, and in particular in statistical physics, and he was a pioneer in developing some of the first statistical models of the analysis of data, such as randomisation. And although he could not have anticipated the quantum revolution, this is an understanding that fits later developments in physics as well.

The four elements that I have emphasised in Peirce's conception of propensity ('objectual', 'causal', 'hypothetical', 'long run') agree to some degree with the maxims of pragmatism. More specifically, maxim two is in line with the 'objectual' character of Peirce's would-be's, maxim four with their 'causal' nature, and maxim five with their being 'hypothetical'. Maxim three resonates with all the features of Peirce's

[23] See for instance the debate with Dr. Carus in Charles Peirce, "Reply to the Necessitarians", op. cit. (1893 / 1935), appendix A, particularly pp. 414-16. Peirce is generally a fierce critic of many aspects of Hume's empiricism but the tone of his criticism is most poignant when addressing Hume's scepticism towards the reality of chances and causality.

dispositional probabilities. Peirce was initially led to the frequency view of probability, and afterwards to the 'long-run' version of the propensity view by pragmatist intuitions. [24] In most respects, his developing views on chance reflect his commitments in pragmatism, and vice-versa: his pragmatism developed in response to his commitment to, and acquaintance with, the reality of chance. Nevertheless Peirce lacked a full understanding of the difference between a statistical feature of an ensemble and the genuinely stochastic behaviour of an individual system. He did not benefit from a fully developed stochastic dynamical theory for any of the phenomena that he was acquainted with. He thus mistook the commitment to a 'long-run' version for a commitment to chance *simpliciter*, and he could hardly have anticipated this element in his theory to be furthest from the pragmatist maxims, including maxim one.

For suppose that our concept of propensity is exhausted by its conceivable effects (including the 'practical consequences' of these effects), as maxim one requires. And suppose moreover that our concept is 'long-run': X has propensity P in experiment E if and only if were E repeated an infinite number of times, the relative frequency of X in the infinite sequence of outcomes of E so generated would be P. What are the practical consequences, or effects, of an ascription of propensity such as this? I already noted that their practical consequences are hypothetical – they are what we would be able to infer from the result of repeating the experiment an infinite number of times. So, on this combination of views, a probability ascription to a single case (to an experiment that can only be performed once) has no meaning. For the effects of the ascription are as a matter of principle not available to us, in any short-enough run, and we are not able to

[24] Arthur Burks (in his "Peirce's Two Theories of Probability", in Edward C. Moore and Richard S. Robin (eds.), *Studies in the Philosophy of Charles S. Peirce* (Armherst: University of Massachusetts Press, 1964), pp. 141-150) appropriately links Peirce's development of propensities to the pragmatist maxim.

conceive them. Since no statement of which we cannot conceive its effects has meaning, the statement that a particular event x in isolation has probability P also has no meaning. There is on his view of chance no coherent concept of single case probability.

Yet, single-case probability ascriptions are not meaningless, and may well have practical consequences, even dramatic consequences – not only in science but in everyday life. [25] This entails that one among our commitments is mistaken, and indeed Peirce seems to have struggled with the tension throughout his life. He was led to a succession of modifications of the pragmatist maxim one with an aim to develop an analysis of practical consequence that would make sense in the long run. But his defence of the indefinite nature of our social commitments within ideally extended communities is both vague and unfeasible for a complete understanding of probability. His aims would have been better served by a single case version of the propensity theory, which would have resolved the tension, since it is simpler to stick to maxim one by relinquishing the commitment to the 'long run'. [26]

Recall what the single-case version entails: Propensities do not manifest themselves as frequencies in infinite virtual sequences, but as probabilities in every single

[25] Peirce was acutely aware of the problem himself: "Yet, if a man had to choose between drawing a card from a pack containing twenty five red cards and a black one, or from a pack containing twenty-five black cards and a red one, and if the drawing of a red card were destined to transport him to eternal felicity, and that of a black one to consign him to everlasting woe, it would be folly to deny that he ought to prefer the pack containing the larger proportion of red cards, although from the nature of the risk, it could not be repeated. It is not easy to reconcile this with our analysis of the conception of chance". Charles Peirce, "The Doctrine of Chances", op. cit. (1878 / 1986), p. 282.

[26] The diagnosis is not new – and neither is the prescription for pragmatism to abandon the long-run view – see particularly James Fetzer, "Peirce and Propensities", in E. C. Moore (ed.), *Charles S. Peirce and the Philosophy of Science* (Tuscaloosa, Alabama: The University of Alabama Press, 1993), pp. 60-71.

experimental trial. This demands a tripartite distinction between the dispositional property of the system (the propensity), its manifestation or effect in each single trial (the probability) and the consequences in the long run of the experiment (the frequency). It is certainly true that the probability that is displayed in every experimental trial is not observable or verifiable in the sense required by empiricists – but then neither propensities nor frequencies are observable or verifiable in this sense on any propensity view, including Peirce's. Yet there is a considerable gain in explanatory power. The long-run view can not bring propensities to bear explanatorily on the single case – no result of a single experimental trial is ever 'explained' in its own terms. The single-case view, by contrast, explains every single trial as the exercise of the underlying disposition, which displays itself in a probability. This explanatory advantage reveals itself fully in the explanation of the sort of genuinely and irreducibly indeterministic phenomena that Peirce presciently understood to be universal – and which is best exemplified by atomic phenomena, such as radioactivity. On a long-run view nothing explains the actual decay of a particular piece of radioactive material – we may only explain the hypothetical decay of an infinite number of similar samples of the material under similar conditions. On the single-case view developed here the actual decay observed is adequately explained by the propensity invoked (the element's 'half-life') and its display in the appropriate probability of decay within a given period of time.

To summarize, the conception of propensity defended in this paper is i) objectual, ii) causal, iii) hypothetical, and iv) single case. It differs maximally from Popper's, which is arguably none of these, and minimally from Peirce's, which is definitely i)-iii) but not iv). It is moreover arguable that the move from a long-run view to a single-case view

can be justified by the very pragmatism that led Peirce to the development of his original view in the first place. And, in addition, the conception defended here rejects a central thesis of later conceptions, including Popper's, namely, the identity thesis. These are all powerful reasons why the label 'pragmatist' is most apt for the conception of propensity defended in this paper.

5. Humphreys' Paradox

The most formidable objection to the propensity interpretation is due to Paul Humphreys and was first recorded in print in 1979 by Wesley Salmon. [27] After a discussion of the relative merits of Reichenbach's frequency interpretation when applied to causal sequences, he writes:

> "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 ¾. It would be strange, under these circumstances, to say that this corpse has a propensity (tendency?) of

[27] In his long review of D. H. Mellor's first book: Wesley Salmon, "Propensities: A Discussion Review", *Erkenntnis*, 14, (1979), pp. 183-216. I will follow the entrenched use and denote the objection as "HP" for Humphreys' Paradox. The term is arguably inappropriate – it is not so much of a logical paradox as a powerful argument against propensity analyses of probability.

¾ to have had its skull perforated by a bullet. Propensity […] seems to inherit the temporal asymmetry of causation." [28]

At this point it may help to distinguish clearly between two different commitments that may be run together in the statement of the identity thesis. There is first the commitment to interpreting probabilities as propensities (the propensity interpretation of probability proper); then there is the commitment to understanding propensities as probabilities. I have so far been referring to the latter on its own as the identity thesis, and we may refer to it as Identity_1. But the full content of the thesis is in fact made up by the conjunction of both commitments. Let us refer to such conjunction as Identity_2. It follows that the identity thesis (Identity_2) entails Identity_1 but not vice-versa.

Suppose first that propensities are (objective) probabilities, in accordance with Identity_1. We may refer to this statement as the propensity-to-probability half of the identity thesis. Propensities may then be written as *Pr* and distinguished from the larger class of all probabilities *P*. The question for the full identity thesis (Identity_2) is then to elucidate for any given probability *P (x)*, whether there is a propensity *Pr (x)* such that *P (x) = Pr (x)*, where the equality sign is crucially taken to denote extensional and not merely numerical or quantitative identity. Let us refer to this as the probability-to-propensity half of the identity thesis. It is the extensional identity that provides the full content of the identity thesis (Identity_2), and it is now clear that there are two ways in which it can fail. It fails if its propensity-to-probability half (Identity_1) fails. But it also fails if its probability-to-propensity half fails.

[28] Wesley Salmon, op. cit. (1979), pp. 213-214.

Note that this is at any rate not how the question is usually framed. I have applied the identity thesis here to absolute probabilities, P (x), and absolute propensities, Pr (x), when it is typically applied to conditional probabilities and conditional propensities. Given any conditional probability P (x / y): is there a corresponding propensity Pr (x / y) that it is identical to, and vice-versa? This way of framing the question takes care of the fact that many philosophers think that all propensities are conditional, [29] while others think that all probabilities are conditional. [30] So I will go along with this assumption for the time being. Let us refer to it as *assumption one*: Propensities are (identical with) conditional probabilities.

Different versions of the HP argument address different versions of the identity thesis. Thus the simplest examples typically deny the probability-to-propensity half; that is, they deny that all probabilities are propensities. Let us for instance consider Salmon's example. We first assume that a particular conditional probability, such as that described in Salmon's example above, is physical and does receive a propensity interpretation. We may then denote by *s* the event type of being shot and by *d* the event type of dying. The probability P (d / s) is then identical to the corresponding propensity Pr (d / s), in the sense that there is a propensity interpretation of such a probability – which roughly coincides with the capacity that shooting has to kill. Let us refer to this as the shooting-to-dying capacity. Humphreys' paradox then arises when we realise that any conditional probability has a well-defined inverse probability, given by the application of Bayes' theorem. If the probability P (d / s) is well defined then so is the probability P (s / d). Yet this inverse conditional probability may in no way receive a

[29] Donald Gillies, op. cit. (2000b), p. 131-2.
[30] Alan Hajek, "What Conditional Probability Could Not Be", *Synthese*, 137 (2003), pp. 273-323.

propensity interpretation, because, as Salmon notes, there is no corresponding dying-to-shooting capacity. So $Identity_2$ is false because the last stage in its demonstration above fails. The reason for the asymmetry lies with the asymmetry of causation; the events are not appropriately listed as cause and effect. While shooting may cause death, it makes no sense to suppose that, in the example provided, dying is a cause of shooting.

It must be emphasised that the problem raised by this example does not *per se* depend on the temporal asymmetry between 'shooting' on the one hand, and 'being shot' – ergo dying – on the other. The temporal and causal asymmetries coincide in Salmon's example, but they need not coincide in general. And what prevents the propensity interpretation of an inverse conditional probability is the inexistence of a causal relation in the direction from the conditioning event to the conditioned event, not any failure in the appropriate temporal order. Another example may make this point more forcefully. Some friends have remarked on my propensity to fly to North America in the spring. Let us denote by 'F' the event of my flying to North America, and by 'S' the event of spring in North America. Both events are extended in time, but 'F' is only a few hours long, while 'S' is three months long. On account of my travelling record over the past 10 years, we may estimate the probability P (F / S) = 0.9, in line with the past relative frequency. Let us then apply Bayes' theorem: P (S / F) = P (F / S) x P (S) / P (F). The prior probability of spring in North America may be calculated by dividing the year into four seasons, which yields P (S) = 0.25. On the basis of such a division and the past relative frequency of my flights to North America, I can also estimate that P (F) = 0.4. We can then work out the inverse probability as P (S / F) = 0.9 x 0.25 / 0.4 = 0.56. This is a well-defined probability, but it has no propensity interpretation. The reason is not any temporal asymmetry of the events, since spring is an extended event in time that

can not properly be said to occur *before*, or for that matter *after*, my flying to North America. The reason is rather that while it makes perfect sense to suppose that Spring in north America is one of the factors that cause my decision to fly there, the converse would be absurd: my flying to North America does not cause spring there. The reason why P (S / T) has no propensity interpretation does not have to do with the time order of the conditioning and conditioned events, which in the example just described is indeterminate, but is related instead to the causal relation between the events – which is certainly determinate in this case.

Thus the propensity interpretation may apply to conditional probability P (y / x) only if the conditioning event x is a cause of the conditioned event y. By contrast, whenever it is absurd to suppose that x may be a cause of y then P (y / x) has no propensity interpretation – regardless of the time order of *x*, *y*. Since causality is typically asymmetric, *y* will rarely be among the causes of *x* whenever *x* is a cause of *y* – again, regardless of their temporal order. Yet Bayes' theorem guarantees that if P (y / x) is well defined then P (x / y) is also well defined (as long as P (x) ≠ 0). In other words, HP shows that there are well-defined conditional probabilities that cannot possibly receive a propensity interpretation. Identity_2 is therefore false (and nothing so far suggests Identity_1 to be false in addition).

I have dealt so far with two very simple but intuitive illustrations, one originally due to Salmon. Humphreys' own discussion appeared in print a few years later, [31] and appeals to a more sophisticated and complex setup. He considers a source that emits photons

[31] Paul Humphreys, "Why Propensities can not be Probabilities", *The Philosophical Review*, 94 (1985), pp. 557-570. See also Paul Humphreys, "Some Considerations on Conditional Chances", *British Journal for the Philosophy of Science*, 55 (2004), pp. 667-680.

spontaneously at some time t_1; a few among these photons reach a half-silver mirror at a certain distance at some later time t_2; some may then be absorbed, but most make it through and are transmitted at time t_3. Let us now, for each photon, denote by B_{t1} the background conditions at the time t_1 of its emission; by I_{t2} the event of its incidence upon the mirror, and by T_{t3} the event of its transmission through the mirror. Moreover, let us suppose, in line with the identity thesis, that propensities are conditional probabilities. Humphreys then argues that the physical situation described dictates the following values for the propensities at the time of emission: [32]

i) $Pr_{t1} (T_{t3} / I_{t2} \,\&\, B_{t1}) = p > 0.$

ii) $1 > Pr_{t1} (I_{t2} / B_{t1}) = q > 0.$

iii) $Pr_{t1} (T_{t3} / \sim I_{t2} \,\&\, B_{t1}) = 0.$

From the perspective of a single-case propensity theory these statements mean the following. Given the background conditions and the emission of a photon at t_1, the photon has some non-zero propensity to reach the mirror. If the photon reaches the mirror, it then has some propensity to be transmitted. Finally, an emitted photon that fails to reach the mirror has no propensity (i.e. has propensity zero) to be transmitted.

Now, following Humphreys, [33] consider the following principle of **conditional independence**:

$$Pr_{t1} (I_{t2} / T_{t3}\, B_{t1}) = Pr_{t1} (I_{t2} / \sim T_{t3}\, B_{t1}) = Pr_{t1} (I_{t2} / B_{t1}). \qquad \text{(CI)}$$

[32] In Paul Humphreys, op. cit. (2004), p. 669. Humphreys' original example (op. cit., 1985, p. 561) assumes that the emission event occurs at an even earlier time, t_0. But nothing substantial in the argument in the text will depend on the assumption $t_0 \neq t_1$.

[33] *Ibid*, p. 561; "Some Condierations on Conditional Chances", p. 669.

CI asserts that the propensity of incidence (I_{t2}) is independent from transmission at (the later) time t_3 given the background conditions at (the earlier) time t_1. CI seems intuitive, but it is unclear exactly why. In particular note that if applied to probabilities rather than propensities, CI would just express a screening-off condition, which is notoriously often inappropriate when applied to the stages of a probabilistic causal process. Given the background conditions, B_{t1}, why should the probability of incidence I_{t2} be independent of T_{t3}, i.e. of whether the photon is transmitted? On the contrary, it seems that a correlation should naturally arise, because the background conditions include also those causal factors that make transmission without incidence physically impossible in the experiment described – an important issue that I shall return to in section 6.

At any rate, it can be shown that CI, (i), (ii) and (iii) are inconsistent with the Kolmogorov axioms of probability. [34] And this entails that the ascriptions of propensity values in the experiment described are inconsistent as long as they are required to agree with both CI and the probability calculus. The outcome is that not all propensities are probabilities, the propensity-to-probability half of the Identity thesis fails, and Identity_1 is shown to be false as well.

The key to the example is CI – why does it seem so intuitive in this case? Humphreys endorses the principle fully, and he states that "the conditional independence principle CI […] claims that *any event* that is in the future of $I_{t'}$ leaves the propensity of $I_{t'}$

[34] Humphreys (op. cit., 1985), p. 562. The inconsistency with Bayes theorem is particularly easy to demonstrate. First, calculate the value of $Pr_{t1}(I_{t2} / T_{t3}\, B_{t1})$ by means of (i), (ii), (iii) and Bayes theorem. We obtain: $Pr_{t1}(I_{t2} / T_{t3}\, B_{t1}) = Pr_{t1}(T_{t3} / I_{t2}\, B_{t1}) \cdot Pr_{t1}(I_{t2} / B_{t1}) / Pr_{t1}(T_{t3} / B_{t1}) = p \cdot q / p \cdot q = 1$. Then calculate it by means of (ii) and CI: $Pr_{t1}(I_{t2} / T_{t3}\, B_{t1}) = Pr_{t1}(I_{t2} / B_{t1}) = q$. Since by definition $q < 1$, we have a contradiction.

unchanged; i.e. $Pr_t (I_{t'} / T_{t''}) = Pr_t (I_{t'})$ [where $t < t' < t''$]. This principle reflects the idea that there exists a non-zero propensity at t for $I_{t'}$ to occur, and that this propensity value is unaffected by anything that occurs later than $I_{t'}$." [35]

However, on the face of it, this is not what CI actually states. Humphreys is rather expressing a more general statement. The idea that is 'reflected' in CI is the more general one that no propensity at time t is affected by anything that happens after t. CI is actually a particular instance of this general idea; it is what results when the general idea is applied to the particular example under consideration, and in particular to the transmission event T_{t3}. Since T_{t3} lies in the future of I_{t2}, the general idea dictates that it can not affect the propensity of I_{t2}, and this is what CI is formally aiming to express. In other words, in referring to CI as the 'principle of conditional independence', we seem to be committing the fallacy of taking the part for the whole, or the instance for its generalisation. The fallacy is innocuous only because what makes CI intuitive is precisely the general idea that it is an instance of. So what then is this undeniably intuitive general idea? Why should the propensity of an event at time t be 'unaffected by anything that occurs later'?

The discussion in the previous section made it clear that the causal asymmetry between a cause *a* and its effect *b* – not their temporal order – makes *P (b / a)* a candidate for a (non-zero) propensity. Hence the general idea underlying CI is intuitive under the presupposition that causes only act forwards in time. It then follows that no events in the future of t' may causally affect anything that goes on at t', so the propensity of $I_{t'}$ can not be affected by anything that happens at t'', where $t'' > t'$. But notice that there is

[35] Paul Humphreys, op. cit. (2004), p. 670, my italics.

nothing peculiar about dispositional properties or propensities at this point: a similar statement concerning categorical properties would be just as intuitive. For any categorical properties *a* and *b*, if *a* lies in the past of *b*, and causes only act forwards in time, then *P (a / b)* cannot be other than zero if it is to represent a propensity. By contrast, neither statement is intuitive if there is backwards-in-time causation. To see this, just carry out the thought experiment: Reject the assumption that backwards causation is impossible, and you will see that both the general idea and its instance in CI cease to be intuitive. If T_{t3} can be considered a backwards-in-time cause of I_{t2} then there is no reason whatever why CI should hold for propensities.

Thus the intuition behind Humphreys' principle of conditional independence (CI) really is just the presumed time-asymmetry of causation. We may refer to it as *assumption two*: The conditioning event of a propensity is a (forwards-in-time) cause of the conditioned event. The assumption is indeed at the heart of HP, and plays the same key role in the simplest examples as well. In Salmon's shoot-to-kill example, the conditional probability P (d / s) may express the conditional propensity of shooting to killing. But the inverse conditional probability P (s / d) does not express a corresponding propensity of killing to shooting, because *d* is no cause of *s*. In the example of my own propensity to travel to North America in spring, P (S / F) does not express an inverse conditional propensity because F is no cause of S. In both cases it is the violation of *assumption two* that gives rise to the air of paradox – the very assumption that we have shown to be involved in the CI principle employed by Humphreys.

6. Some responses to Humphreys' Paradox

The identity thesis is essentially involved in both assumptions one and two, and in section 8 I will suggest that the appropriate response to HP is to reject the thesis altogether (and, concomitantly, to reject assumptions one and two). In this section I review some of the main alternatives that have been proposed in response to the problem raised by HP over the last two decades, and I argue that they implicitly or explicitly commit to the identity thesis.

Critical responses to HP in the literature may be classified in two distinct types, depending on whether they reject the HP argument as invalid, or accept the argument but disagree with Humphreys' interpretation of its conclusion. Note that regardless of what class they fall into, all critical commentators accept that there is a genuine contradiction in the example provided between CI and the axioms of classical probability theory. The differences rather concern the significance of this contradiction – and consequently, the actual conclusion of the HP argument.

Among the first type of responses, McCurdy attempts to show that the HP argument is invalid because the example considered does not bear out the principle of conditional independence, which he believes to be false in the example given. [36] Indeed it was noted in the previous section that CI is likely false if read as a mere statistical screening-off condition (note the replacement of propensity functions *Pr* with probability functions *P* in the following expression):

$$P_{t1}(I_{t2} / T_{t3}\, B_{t1}) \neq P_{t1}(I_{t2} / \sim T_{t3}\, B_{t1}) \neq P_{t1}(I_{t2} / B_{t1}). \qquad \text{(CPD)}$$

[36] Christopher McCurdy, "Humphreys' Paradox and the Interpretation of Inverse Conditional Probabilities", *Synthese*, 108, (1996), pp. 105-125.

We may refer to this inequality as the conditional probabilistic dependence (CPD) between the events cited. These events may themselves have many causes, which could explain their probabilistic dependence. For instance, McCurdy himself points out that there may be common causal factors in the past causal history of both T_{t3} and I_{t2}. There could also be a residual correlation between the exogenous variables that give rise to T_{t3} and I_{t2}. Or there may be 'brute' probabilistic dependence if the causal processes that give rise to T_{t3} and I_{t2} are genuinely indeterministic.

However, Humphreys meant conditional independence (CI) to be something other than screening-off; otherwise he would not have used propensity functions, Pr, in the definition of this condition. As the general principle that gives rise to CI makes clear, CI is instead meant to express a form of causal independence characteristic of propensities. And CI is true if understood in this way as a *sui generis* form of causal independence, as long as *assumptions one* and *two* are maintained. It then makes no sense, in the experiment described, to suppose that the propensity of B_{t1} to produce I_{t2} depends on whether or not T_{t3} occurs. (The reason, as already mentioned, is not the temporal order of I_{t2} and T_{t3} per se, but the asymmetry built into the causal process). [37]

Most other authors have endorsed the HP argument and claimed it to be a major obstacle, if not definite indictment, for any propensity analysis of probability. The differences between these authors are mainly in emphasis and reflect further background commitments. Thus Salmon argues that HP should lead us to abandon the

[37] Nevertheless, as noted already, this form of causal independence carries an implicit commitment to those two assumptions, only one of which is causal in nature, namely assumption two. Assumption one, by contrast, is representational: it provides a formal representation for conditional propensities in the calculus of probability as conditional probabilities. Thus we could say that CI is two-faced: it represents both a screening-off relation and a characteristic causal relation between propensities and their effects.

propensity interpretation and embrace instead the frequency interpretation that he favors. [38] His argument may be reconstructed as follows: i) there is objective probability, or chance; ii) chance requires an interpretation; iii) the only possible interpretations of chance are frequencies and propensities; iv) HP shows that the propensity interpretation is flawed; hence v) chances ought to be interpreted as frequencies. Most commentators see this as an unnecessarily harsh reaction. [39] But while other commentators have mainly disputed assumptions i) or iii) in Salmon's argument, I believe that the most problematic assumption is ii). The reason will become clearer in the next few sections, but let me briefly anticipate its implications for Salmon's argument. Suppose iii) is true and the objective probabilities can only be frequencies or propensities. If so, supposing ii) in addition is tantamount to accepting an expanded version of the identity thesis ($Identity_1$), which identifies probabilities with *either* propensities or frequencies. Yet, each of these identifications is flawed or at least greatly problematic. [40] Since i) is a prerequisite for any philosophical theory of chance, we must reject ii) instead: chance (understood as objective probability) needs no interpretation. [41]

---

[38] Wesley Salmon, op. cit. (1979).

[39] For reactions to Salmon's argument see, for examples, Antony Eagle, "Twenty- one Arguments against Propensitiy Analysis of Probability", *Erkenntnis*, 60 (2004), p. 407; Donald Gillies, "Varieties of Propensity", *British Journal for the Philosophy of Science* (2000a), section 6; and Peter Milne, "Can there be a Realist Single-case Interpretation of Probability?", *Erkenntnis*, 25 (1986), pp. 129-132.

[40] This paper reviews the case against the identity thesis with respect to propensities. For arguments against what is arguably essentially the same thesis but with respect to frequencies see e.g. Alan Hájek, "Mises Redux' – Redux: Fifteen arguments Against Finite Frequentism", *Erkenntnis*, 45 (1997), pp. 209-227; "Fifteen Arguments Against Hypothetical Frequentism", *Erkenntnis*, 70 (2009), pp. 211-235; Donald Gillies, op. cit (2000b), chapter 5); and D. H. Mellor, *Probability: A Philosophical Introduction* (London: Routledge, 2005), chapter 3.

[41] A referee directed me to a possible antecedent in Isaac Levi's view that chance requires no semantics in *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance* (Cambridge, Massachusetts: The MIT Press, 1980) chapter 12.

Milne's response is similar to Salmon's, endorsing HP as a fundamental objection to propensities. [42] But he goes further than Salmon in claiming that HP shows that the very idea of single-case propensity is flawed or incoherent. At first sight this argument may appear not to rely on the identity thesis at all. In particular, it appears not to require that all (objective) probabilities are propensities. However, it does presuppose the other half of the identity thesis, namely, the propensity-to-probability half, since it assumes that all propensities are (conditional) probabilities. (In the terminology of this paper, Milne's argument assumes Identity_1).

Finally, Humphreys own reaction is in my view the most judicious one. He does not think HP gives grounds to reject the concept of propensity, nor does he think it constitutes an argument in favour of the frequency interpretation. Instead he claims that HP suggests the need to abandon the classical probability-calculus representation of propensities as (classical, or Kolmogorov) probabilities, and then asks: "If conditional propensities cannot be correctly represented by standard probability theory, what does that say about the status of probability theory?" [43] It is unclear exactly how else Humphreys intends to represent propensities. He is certainly committed to the identity thesis in the derivation of HP, and his response seems to suggest that an alternative calculus of probability be developed precisely in order to preserve the identity thesis.

7. Pragmatism and Humphreys' Paradox

[42] Peter Milne, "Can there be a Realist Single-case Interpretation of Probability?", *Erkenntnis*, 25 (1986), pp. 129-132.
[43] Paul Humphreys, op. cit. (2004), p. 679.

The application of the identity thesis to Humphreys' example is one-way: all propensities are understood as probabilities, but it is not required in addition that all probabilities be interpreted as propensities. The exceptions may come in three varieties. First, the identity thesis already leaves it open that in addition to the physical chances that propensities aim to interpret, there may exist personal probabilities or credences. Accounting for such probabilities may require a different interpretation – such as the subjective or Bayesian interpretation. Second, ever since Kolmogorov wrote, there is an entrenched distinction in the literature between absolute and conditional probabilities, and it may well be that the propensity interpretation fits one but not the other kind. In particular perhaps there are no *absolute propensities*, and only conditional probabilities are interpretable as propensities. [44] Finally, the conclusion of the HP argument makes it clear that if a conditional probability has a propensity interpretation, then its inverse – which is also a conditional probability – will most likely not have it. So the probability-to-propensity half of the identity thesis seems incompatible with the assumptions built into the HP argument.

However, the propensity-to-probability half of the identity thesis (Identity_1) is presupposed in the HP argument, where it is introduced by means of assumptions *one* and *two*. Both assumptions are the offspring of the empiricist instinct to reduce problematic theoretical concepts to empirical ones. *Assumption one* aims to reduce propensity to probability, while *assumption two* aims to reduce it to causation. *Assumption two* in addition implicitly rules out backwards-in-time causation, which has been an important part of empiricist thinking ever since Reichenbach's influential

[44] This is the view defended by Gillies, op. cit. (2000b), pp. 131-132.

attempt to reduce the direction of time to open conjunctive forks. [45] In other words, both assumptions make full sense in the context of successive empiricist attempts to reduce time and propensities to causation, and causation to statistical correlation. Such reductionisms have on the whole been alien to the tradition of pragmatism. When they have cared for such issues, pragmatists seem to have been happy to accept that the direction of time may be primitive, that backwards-in-time causation is conceptually possible, and that a cause may leave the probability of its effects unchanged. Certainly, in terms of the version of pragmatism reviewed in section 2 of this essay, maxims two and four explicitly oppose such reductionism, and seem in open conflict with both assumptions.

Let us then consider the possibility that both assumptions are mistaken. Suppose that contrary to *assumption one* propensities are not to be identified with probabilities, and in particular may not be appropriately represented as conditional probabilities. But what else could propensities be, and how else could they be represented? Here a change of paradigm, from an empiricist to a pragmatist framework, may be helpful. Suppose following Peirce that propensities are theoretical properties ascribed to objects by scientists in an attempt to explain phenomena involving those objects. Propensities are on this view not to be identified with probabilities. Instead they are more generally taken to be dispositional properties with probabilistic displays or manifestations. There is on this view no need to represent the relation between the propensity and its manifestations as a conditional probability, or any other probability for that matter. An appropriate representational schema would introduce a new symbol » to represent what is after all a new and *sui generis* relation. Let us first consider the case of a deterministic

[45] Hans Reichenbach, *The Direction of Time* (Los Angeles: University of California Press, 1954).

dispositional property. We may then write D » M for "D manifests itself as M". For example, the fragility of a glass may manifest itself in its breakage, so we may write: Fragility » Breakage to represent the particular manifestation event, and to distinguish it from a breakage that does not display fragility.

A dispositional property is a propensity if it possesses more than one manifestation property with some probability. The manifestation event then consists in the display of a probability distribution over these properties: D » P (M_i), where M_i are the possible manifestation properties. (Alternatively: the manifestation property of a propensity has several values, and the propensity displays itself in a probability distribution over these values). Thus tossing a fair die (D) has six manifestation properties (M_1, M_2, … M_6), and the 'would-be' of the die displays itself in a probability distribution over these values: D » P (M_i) = 1/6.

Thus we can see that the rejection of *assumption one* already implies a rejection of the identity thesis, in particular its propensity-to-probability half. What about *assumption two*? The pragmatist need not reject upfront that the relation between a propensity and its manifestation properties is causal, but can keep a neutral stance instead. It is arguably an empirical matter to determine, for any particular propensity, whether it is causally related to its manifestations. Similarly for backwards causation: it is an empirical matter whether some causes act backwards in time. Hence *assumption two* is no longer required, for three reasons. First, there is no longer an a priori commitment to the causal nature of the relation between the propensity and its manifestation. Second, there is no commitment to causes preceding their effects. Finally, since we have already disposed

of *assumption one*, there is no commitment to the conditional-probability representation of effects as events conditioned upon their causes.

Let us now return to Humphreys' Paradox. The application of the pragmatist conception of propensity to the HP argument requires a new representation of Humphreys' experiment and, in particular, of those conditions i-iii) employed to characterise the propensities involved. These propensities are best represented by means of the following conditions in the new notation instead:

a) B_{t1} & I_{t2} » P (T_{t3}) = p, where p > 0.

b) B_{t1} » P (I_{t2}) = q, where 1 > q > 0.

c) B_{t1} & ~ I_{t2} » P (T_{t3}) = 0.

Conditions a-c) represent the appropriate propensities of the experiment in terms of the events described without any recourse to conditional probability. On the dispositional view defended here, these are propensities of the systems involved, perhaps even the entire experimental arrangement, at some stage in their temporal evolution. [46] They are not strictly speaking properties of the events described on the left hand side of the symbol ». But since an event is typically understood as the change in properties of some system, we may consider the whole system or systems involved in such a change. We may then use the events as shorthand for the description of the dispositional properties of the system up to the time of the last event mentioned. So the left hand of the

[46] David Miller (in *Critical Rationalism: A Restatement and Defence* (Chicago: Open Court, 1994), pp. 183-84) urges that any propensity at time *t* is implicitly conditional on the state of the whole universe at *t* and so the relevant system turns out to be the whole universe. The formulation above is deliberately vague in this regard; different views follow from different specifications of what counts as the "system". Note that this is compatible with an 'objectual' as opposed to 'relational' conception of propensity.

expressions a), b) and c) describe some change in the properties of the systems that are endowed with propensities – and in particular in the particles themselves as they move along their trajectories. The right hand side of the expressions describe the manifestations or displays of such propensities as probability distributions over the values of the relevant manifestation property.

The new representation solves the problem that gave rise to Humphreys' paradox straightaway. There is no contradiction between CI, Bayes' theorem, and conditions a)-c) above. The reason is simple, and quite independent of whether or not CI holds. The application of CI and Bayes' theorem to conditions a) – c) has no effect, since the probabilities that appear in the right hand sides of those conditions are not conditional probabilities. In other words, the problem is shown to be an artefact of the conditional-probability representation of propensities. [47]

8. A Comparison with the Empiricist Alternative

In this section I consider the most outstanding 'empiricist' alternative to the 'pragmatist' conception of propensity defended in this paper, namely, Donald Gillies' long-run version of the propensity theory. I argue that this theory also fails to solve Humphrey's paradox since it is ultimately committed to the identity thesis.

[47] There is a sense in which this is line with Humphreys' own conclusion that propensities are ill served by the classical probability-calculus representation (see the end of section 6). But Humphreys suggests in response to adopt a different probability calculus in order to represent propensities, while I am suggesting that propensities cannot be uniformly represented as probabilities, under any calculus, precisely because they are distinct from their probabilistic manifestations.

Gillies' theory is in many respects very close to Popper's original 'earlier' views. It is a long-run theory that takes propensity ascriptions to be testable, just like any statistical hypothesis. For Gillies, like Popper, the criterion of testability for propensities is essential – thus squarely placing them both in what I have called the "empiricist" tradition. In particular what Gillies calls a falsifying rule for probability statements (FRPS) is a method to fix a significance level for any statistical hypothesis, which thus provides a means to test it, as follows. Consider a statistical hypothesis H and suppose we are trying to test H against some evidence, consisting in a sample of n data points (e_1, e_2,…., e_n). Let X be a test statistic, that is to say a function X (e_1, e_2,…, e_n) of the observed data whose value can be calculated from the data. Suppose that we can derive from the hypothesis H that X ought to have a particularly shaped distribution P (X) – call any such P (X) a 'falsifying distribution'. We can then separate a head and tails part to the distribution, as follows: $a \leq X \leq b$. They are such that the probability of obtaining a result in the tails is below the significance level, typically set at 5%. Now the falsifying rule for probability statements states that "if the value obtained for X is in the tails of the distribution, this should be regarded as falsifying H; whereas, if the value of X is in the head of the distribution, this should be regarded as corroborating H". [48] It is obvious that this is not strict falsificationism, but a pragmatic form of methodological falsificationism instead.

So far Gillies' theory is simply an empiricist version of a hypothetico-deductive theory of confirmation applied to statistical hypotheses. These may be represented by means of probability spaces. A probability space is an ordered triple (Ω, F, P), where Ω is the sample space, F is the Borel field of subsets of Ω and P is a real-valued function defined

[48] Donald Gillies, *Philosophical Theories of Probability*, p, 147.

on F. These encapsulate Kolmogorov's axioms, which are expressed by Gillies in terms of a single Axiom I: P is a non-negative, countably additive set function on F such that P (Ω) = 1.[49] Thus in stating that "in the present version of the propensity theory the link is established instead by adopting the falsifying rule", Gillies is in effect committing himself to the identity thesis, because the falsifying rule links statistical hypotheses (probabilities) to test statistics over the data (frequencies).

But the link between probabilities and frequencies is as a matter of fact not so straightforward. It requires an additional axiom, referred to by Gillies as *the axiom of independent repetitions*, [50] which is worth discussing in some detail since it requires an expansion of the usual notion of probability space. It first introduces a fourth element sui generis for propensities, namely, a *sequence of repetitions* S_s, which is defined as follows. A sequence of events (such as the outcomes of each particular experimental trial) is a sequence of repetitions relative to a set of conditions S_s (which include a spacing condition *s* that regularly separates such events in space or time), if all the conditions *S* are satisfied by each event and the events are separated as required by *s*. Gillies then introduces the notion of a probability system, which extends a probability space to include sequences of repetitions as follows. (S_s, Ω, F, P) is a probability system if (Ω, F, P) is an ordinary probability space and Ω is the set of possible outcomes of the repeatable conditions S_s. Suppose that we then select an n-tuple of such repeatable conditions a repeated number of times, i. e. a new sequence of repetitions denoted as S^n_s. We may then construct a new probability system considering the n-fold Cartesian product of Ω, which we may denote as Ω^n. Similarly we let F^n be a Borel field of subsets of Ω^n defined as the minimum Borel field containing F defined over the

[49] Gillies, *Ibid*, p. 160.
[50] Gillies, *Ibid*, p. 164.

Cartesian products of the elements of *F*. The *Axiom of Independent Repetitions* (axiom 2) then states that: If (S_s, Ω, F, P) is a probability system, so is (S_s, Ω^n, F^n, P^n) for any *n*, where the measure P^n on F^n is the n-fold product measure of the measure *P* on *F*.

Gillies then shows that the empirical laws of probabilities follow from axiom I and axiom II together with the falsifying rule for probability statements (FRPS). Axiom II is interesting in bringing in a further element, the sequences of repetitions S_s. But as far as the identity thesis is concerned their introduction does not change matters, since sequences of repetitions are not themselves dispositional properties, or propensities, distinct from probability distributions. Rather, as we saw, sequences of repetitions are sets of events, and typically they are just the outcomes of a particular kind of experimental trial (given the spacing condition). Recall that probabilities are defined over events or propositions, while propensities (understood as dispositional properties) reside in the objects or their properties. So it does not seem *prima facie* that the introduction of sequences of repetitions allows us to introduce propensities as distinct from probabilities. Thus Gillies' theory lacks the resources to evade a commitment to the identity thesis and, consequently, to overcome the HP objection.

As an alternative, we may choose to focus upon the experimental systems themselves. On Gillies' view it seems that a particular kind of trial either characterises, or is characterised by, the conditions S satisfied by all events in the sequence S_s. The bearers of the propensities can then be the objects that make up the experimental apparatus and at least partly constitute the set of conditions S, or, alternatively, the conditions themselves understood as generalised properties of the whole experimental system. In either of these cases, propensities are most definitely not probabilities, since they are not

borne by those objects that appear in the sigma field over which the relevant probability distributions are defined. Their bearing objects in fact appear neither in the probability space (Ω, F, P) nor in the expanded probability system that encompasses it, namely, (S_s, Ω, F, P). The propensity-to-probability half of the identity thesis would fail on any of these interpretations because "propensities" would not be here identified with any element within either a probability space or a probability system. There is therefore some room to explore ways to get around the HP argument.

However, on reflection, neither of these interpretations can be Gillies' propensity theory. In particular note the following two commitments in Gillies' theory: i) it is a long-run as opposed to a single-case theory, and ii) it is an empiricist theory in the sense that propensity claims are supposed to be methodologically falsifiable. As regards i) if propensities are properties borne by the experimental apparatus, or the whole experimental system, then they are not 'long-run', since they obtain, or display themselves, in every experimental trial. Gillies' theory would be indistinguishable from a standard single-case theory. As regards ii) on these interpretations propensities are not testable in the way that statistical hypotheses are, and the falsifying rule for probability statements (FRPS) strictly speaking does not apply to propensities, but only to the probability distributions that putatively ensue from them. It would be wrong on any of these interpretations to claim that propensities are testable like forces and masses in Newtonian mechanics as Gillies and Popper both claim.

There is yet a third option. In a deflationary spirit we could consider the events that compose the sequence themselves as the bearers of the propensities. So it would be the outcomes of the experimental trials that bear the propensities. Since the probability

distribution P is only defined over Ω, the outcome space, and not over the sequence of repetitions S_s, the propensity-to-probability half of the identity thesis would seem to fail. But this move again confronts the usual objections: the resulting theory is not long run (for each outcome of the experimental trial will presumably display the propensity), and propensity statements would not strictly speaking be falsifiable, but only the statistical hypotheses that they entail. Both problems would presumably be solved by expanding the probability distribution function P^n, while defining it over the Cartesian product of $(S_s \times \Omega^n)$. But this would come at the cost of a reintroduction of the identity thesis, and would not solve the HP objection. Etc.

In summary, there are a number of ways to interpret Gillies' theory. However, none of the interpretations simultaneously agree with all the explicit commitments of the theory *and* force a rejection of the identity thesis. In particular those interpretations that fit in with both the 'long-run' and 'empiricist' commitments explicit in Gillies' theory are also committed to the identity thesis. As a consequence Gillies' empiricist long-run version of the propensity theory cannot provide a solution to the HP objection in anything like the terms explored in this paper.

COMPLUTENSE UNIVERSITY OF MADRID MAURICIO SUÁREZ