Beyond the Magical Thinking Behind the Principal Principle

EDWARD JAMES

Abstract
David Lewis's *Principal Principle* (PP) states that our credence in a single case follows from the general probability of all such cases. Against this stands the Challenge Argument (CA) – to show that the inference is justified. Recent (1) law-to-chance, (2) Bayesian, and (3) propensity theories of probability take up the challenge – but, I argue, fall short. Rather, we should understand (4) propensity via Aristotle's analysis of spontaneity (5) and probabilistic reasoning via the *Anti-PP* and (6) the practice of *bundling one offs*, where (7) *forced bad-odds one offs* illuminate how extensive a role luck plays in our lives.

Our probabilistic reasoning is marred by the magical thinking that probability applies to the single case, a belief that lies deep both in our language itself – as we make such claims as 'the odds are in our favor *in this case';* 'the chances *here and now* are …' – and also in influential and diverse interpretations of probability that each with qualifications accept David Lewis's claim that 'the only basic principle we have about chance, the principle that tells us all we know' is the *Principal Principle* (PP) –

$$C(a_o|e)|P^*(a|e = r) = r$$

– in which our credence, C, or strength of belief in a single case, a_o, given our evidence, e, is, or should be, equal to the general probability, P*, of a as a type and as conditioned by the evidence, e. In short, 'Chance is objective single-case probability' and as such is operative in the single case.[1]

The PP itself represents a rejection of a long-standing view that probability could not be applied to the single case because of its

[1] David Lewis, 'Humean Supervenience Debugged', *Mind*, NS, **103** (1994), Symposium: Chance and Credence, 475 and 483. The various interpretations, as we will see, tone down Lewis's exuberance and then frame the PP in their own terms.

doi:10.1017/S0031819115000121 © The Royal Institute of Philosophy, 2015
First published online 21 April 2015

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

essential tie to observed frequencies.[2] The standard view in effect presented a *Challenge Argument* (CA): that a probability claim is an idealization of a long-run frequency distribution and that hence there is no way, minus a magical jump, to infer, or even give meaning to, the probabilistic property of an item in that distribution, $a_{o,}$ from the property of the distribution itself – P*(a).

Recently, the challenge has been taken up, prompted by a Moorean-type intuition that we legitimately apply probability to the single case;[3] by a belief that the objective probability of quantum mechanics seems to put single-case probability at the forefront, as scientists speak of the probability of the decay of this or that particular particle; and by a belief that the CA could be met.

In this essay I will spend some time considering each of the main views that seek to meet the CA – (1) law-to-chance, (2) Bayesian, and (3) propensity theories – for only by doing so can we feel the grip that the magical thinking of single-case probability reasoning has on us and so free us from an orientation that is not only misleading but also harmful by giving us a false confidence in our single-case reasoning. Then, more positively, I will (4) advance an understanding of propensity that builds on Aristotle's analysis of spontaneity and then argue that probability (5) guides us by the Anti-PP, that warns us not to apply probability to, or to bet on, the single case, (6) directs us to *bundle* never to seldom repeated cases, *one offs*, into long-run projections, and (7) given the many *forced bad-odds one offs* we face, illuminates how extensive a role luck plays in our lives. In the end, though, after Wittgenstein's observation that philosophy leaves everything as

[2] Four decades ago Henry Kyburg, in his seminal *Logical Foundations of Statistical Inference* (Dordrecht: Holland, D. Reidel Publishing Company, 1974) stressed that 'probability can be attributed only to kinds of events or sets and not (usefully) to individual events or to members of the sets that belong to S ... We cannot talk about the probability that John Smith, a 40-year old coal miner will survive for a year. This fact has been emphasized by Neyman, by Reichenbach, by Salmon, by von Mises, by Cramer – by nearly every serious writer who has adopted an empirical interpretation of probability: probability is applicable only to sets of events or kinds of events' (8).

[3] Lewis himself – in his 'A Subjectivist's Guide to Objective Chance', in *Studies in Inductive Logic and Probability*, Vol. 2, ed., R.C. Jeffrey, (Berkeley: University of California Press, 1980) – grants that chance is 'at bottom ... really just a "folk" concept', but one that carries with it an objective truth (269). We shall see this idea – that chance, as objective single-case probability, is basic in human belief – variously affirmed throughout the essay.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

it is, I will have said nothing other than, albeit perhaps illuminated more, what we already implicitly know.

1.

Law-to-chance theories deny the CA by holding that probability purports to provide a law or a law-like claim – an idealization of an observed frequency – and that like any law, if true, holds in the single case, and holds in a way that can only be construed as an entailment. Lewis is very clear here, as he explains that why chance is objective single-case probability 'lies in remembering that single-case chances follow from general probabilistic laws of nature' – or that 'it seems intuitively right to rely on this information – namely, that our credence should follow from our understanding of chance… which itself is determined by a probability law that is projected into the future'.[4] Lewis justifies this deductive relationship on the grounds that chance is objective – supervenient on and so determined by universal laws. Thus Lewis exclaims that 'this is not epistemology! … I am not talking about how evidence determines what's reasonable to believe about laws and chances. Rather, I'm talking about how nature – the Humean arrangement of qualities – determines what's true about the laws and chances. Whether there are any believers living in the lawful and chancy world has nothing to do with it'.[5] Hence, while probability is based on long-run frequency distributions, probability still figures in the single case; for chance, the objective probabilistic property of the single case, 'follows from' or is 'determined by' the objective probability of the sequence itself. Thus, *contra* CA, one can legitimately infer $P(a_o)$ from $P^*(a)$ just because $P^*(a)$ includes $P(a_o)$.

Part of the lure of this reasoning is built on a seeming parallel of probability laws and laws that yield *regular regularities* – regularities that provide a direct link from the idealized regularity, the law, to a particular instance of the law. For example, we observe that objects fall at the rate of $d = \frac{1}{2}gt^2$, that this, a_o, is an object, and so, *ceteris paribus*, deduce that it, a_o, must fall (F) at the rate (R) of $d = \frac{1}{2}gt^2$: (a) $(Fa \rightarrow Ra)$ & Fa_o / thus, Ra_o. So, too, by parallel reasoning, we observe that under cloud cover of type a it rains n % of the time and that this single case, a_o, is a cloud cover of type a, and so, again,

[4] Op.cit. note 1, 477, 485.
[5] Ibid. 481–482.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

deduce that it, a_o, must have an n % chance of raining:

$$(1) \quad [P^*(a) = n \& a_o] \rightarrow C(a_o) = n$$

$$(2) \quad [P^*(a) = n \& a_o] / \text{thus } C(a_o) = n.$$

The reasoning seems so natural, but only in the way that the fallacy of division is natural, as we too readily infer a property of the single case, $C(a_o) = n$, from a property of a collective, the long-run idealized frequency in (1) – $P^*(a) = n$. Laws that characterize regular regularities cover or entail their single cases: any departure of a single case from the law is regarded at best as an anomaly and at worst as a refutation. In sharp contrast, probability laws are *irregular regularities*: they are sequences with multiple results where the results empirically hover around some mathematical idealized ratio – the regularities – but have no algorithm as to what particular event (the single case) happens next – the irregularities. Consequently, such laws, by definition, do not cover or entail the single case at all but only the long-run frequency. This is why no single case can refute any probabilistic claim; if the long-run frequency does not hover around the idealized ratio, then, and only then, may we question the law[6] This is also why we regard it as childish to complain when our 'really good odds' – say, betting on the straight flush – do not come out 'as expected' and so tell that 'child' that such irregularities happen, so deal.

More recently, Jenann Ismael has reworked Lewis's view of chance while keeping to his central idea that chance is probability law applied to the single case.[7] Like Lewis, she sees the PP as basic, as a deep-set

[6] Throughout this essay I will slide by two difficult problems – the reference class problem: how, if at all, to determine just what probabilistic distribution a single case best fits into; and the (dis)confirmation problem: how to (dis)confirm idealized ratios – e.g., to determine how long a sequence must be before it can (dis)confirm some probabilistic projection. Whatever answers are given to these questions will not negate or weaken what I say about the inapplicability of probability to the single case. For any view on applied probability faces these two problems; and what I say throughout this essay about single-case probability is independent of one's solutions to either of these problems.

[7] As Jenann Ismael puts it – in her 'A Modest Proposal about Chance', *The Journal of Philosophy*, **CVII** (2011) – while 'general probabilities apply to classes and do not generally have a time index', '[s]ingle case probabilities … are unconditional, time-dependent, and pertain to particular occurrences' – e.g., 'that *this particular roll* of *these particular dice* … on *this day* comes up double sixes' (421 and 418). (Ismael's emphases)

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

folk belief: an analytic truth anchored in our linguistic framework, one of the 'things that anyone who "grasps the concept" of chance will be able to tell you'.[8] However, in developing her concept of chance, Ismael leaves Lewis's objective orientation toward law and chance for an epistemic and 'pragmatic reorientation ... that physical theories are practical and epistemic tools ... not part of the fundamental ontology ... [but] best thought of as partially prepared solutions to frequently encountered problems' – what 'holds for creatures like *us*'.[9] And what in particular holds for creatures like us is a probabilistic law that is inherently connected to our temporal location – namely,

$$\text{'Def:}Ch_t(e) = Pr_G(e/\text{pre-t history})$$

where $Pr_G(A/B)$ is the general [limit-case, idealized frequency] probability that a random pick from the B's will yield an A' and where $Ch_t(e)$ is 'the chance of *e* assessed at t'.[10] By Def in particular we can escape Lewis's exorbitant understanding of chance as holding for all time: we can now both say that Napoleon had a chance of victory at Waterloo – for the chance then was understood in terms of pre-Waterloo history – and also say that there is no chance now, post Waterloo, that Napoleon could be victorious. The primary advantage of Def by far though is its unifying role: it 'explains the epistemic role of chance' by specifying chance as 'merely a special case of ... Pr_G'.[11]

As impressive as Ismael's advances on Lewis are,[12] to show that chance is a special case of Pr_G she must show how this understanding

[8] Ibid. 420.

[9] Ibid. 442 and 441. Ismael later further frames her pragmatic orientation in terms of a block or timeless universe, thereby taking any dynamic chanciness out of probability. But none of this directly bears on her reply to the CA and can be bypassed.

[10] Ibid. 431 and 425. As Ismael writes: 'The only specific information a situated human agent has about the world is information about past events' (439). Later, she sharpens this definition to make it more in line with relativistic theories by speaking of the light cones of a point, *p*, rather than the history of e at t (436) and by placing law and chance within a relativistic timeless or block universe.

[11] Op. cit., note 7, 430.

[12] Perhaps Ismael's most impressive advance on Lewis is found in her essay, '**Raid**! Dissolving the Big, Bad Bug', *Nous* **42** (2008), 292–307 – in which she dissolves Lewis's worry that a probability law, P*, would admit as probable a competing law, P_1, that would in turn contradict the probabilities of P*. To sidestep Lewis's dodgy appeal to 'inadmissible evidence',

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

meets the CA. How can she infer $Ch_t(e)$ from Pr_G? In reply, Ismael tries to connect Lewis's deductive justification of the CA with another inference that is also implicit in Def, namely, the inference to $Pr_G(e/\text{pre-t history})$ from the evidential basis. Ismael defends this inference, what I will call the *upward link*, by relying on Bernoulli's law: 'as the size of the ensemble increases', it typically approaches the objective probability $Pr_G(A/B)$. This law provides 'a necessary, probabilistic [upward] link between probabilities and frequencies'. However, the law comes with a 'bad-news addendum: … The possibility of divergence remains, no matter how large the ensemble'.[13] That is, the more items we observe 'in a typical ensemble of systems in ψ' the more likely it is, but with no guarantees because of the bad news addendum, that 'the general probability of a/ψ' is approached.[14] And this in turn, Ismael writes, 'gives us' the *downward* 'chance-frequency link', which she expresses formulaically: 'if S is in a state ψ there is a necessary but approximate and probabilistic relation between the chance of observing *e* in an x-measurement on S and the relative frequency of *e* in x-measurements on systems in ψ'.[15]

But first, how the upward link connects with the downward link Ismael never makes clear. All she says, laconically, is that the upward link 'gives us', or will 'provide' us with, the downward 'chance-frequency link'. Yet to observe that some set of items, all these a's together, hover around a/ψ – the basis of the upward link – does not 'provide' us with any downward link – that this a_o hovers or tends to hover around a/ψ – anymore than it is to observe that the average height of a male giraffe is 5.2 meters is to say that the height of this male giraffe here and now is 5.2 meters or even tends to be 5.2 meters. Our knowledge of the properties of an aggregate like a frequency does not rely on the items of the aggregate having the properties of the aggregate.

---

evidence that allows him to rule out competing laws, Ismael distinguished between 'reasoning within the scope of a theory' and 'reasoning about theories' (296), where reasoning within the scope of the theory does not allow for competing probabilistic laws and so does not generate competing probability claims.

[13] Ibid. 421.

[14] Ibid. 422.

[15] Ibid. Ismael initially defines probability in terms of quantum mechanics (419) but throughout the essay uses instances from ordinary life, on behalf of 'creatures like us' (442).

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

And second, whatever one thinks of the upward link as framed by Bernoulli's law, the downward link is a straight-out fallacy. The fallacy can be readily seen by applying Ismael's formulaic presentation of the downward link above to a more concrete instance, say: If the weather [S] is in a particular barometric state [ψ] there is a necessary but approximate and probabilistic relation between the chance of observing rain here and now [e_o] given a barometric reading [an x-measurement] on the weather [S] and the relative frequency of rain [*e*] given barometric readings [x-measurements] on systems in a particular barometric state [ψ]. Clearly, this commits the same fallacy of division found in Lewis: it jumps from the property of the frequency – the relative frequency of e (rain) given x-(barometric) measurements on systems in ψ – to a property of a member of the frequency – the chance of observing e_o (rain) here and now given an x-(barometric) measurement on S (the weather). Whether chance is objective and metaphysically dynamic, after Lewis, or merely epistemic and pragmatic, after Ismael, we still can't cross the chasm from the idealized frequency of the aggregate to the single case.

2.

One quick and easy way to cross the probability-chance chasm would be to reject the CA's assumption that there is a need to legitimately infer $P(a_o)$ from $P^*(a)$ for the reason that probability claims are not based on our observations of any long-run frequencies but on our willingness to bet, to lay odds, on the single case. Hence there is no jump from probability to the single case, for probability, defined by the odds we are willing to lay on the single case, is of the single case. But such a way out offers the hocus pocus of bare belief in place of the surer grounding that observed frequencies supply. As Clark Glymour observed, such credence is 'remote' from ordinary life as well as from science, for it is about one's opinion, not about nature or the legitimacy of one's daily decisions.[16] Probability should guide us and not we it.

Bayesians, like Colin Howson and Peter Urbach, seek to escape this charge of mere subjectivism, while retaining the idea of probability as built on the practice of a bet, through grounding probability in objective mathematical (Bayesian) formulas and then arguing that these formulas both support and then, indeed, require the PP. They in particular work from the first form of

[16] Clark Glymour, 'Instrumental Probability', *Monist* **84** (2001), 285.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

Bayes's Theorem –

$$P^*(h|e) = [P(e|h) \times P(h)/P(e)]$$

– where the right side determines the left. That is, to determine P* – *the posterior probability*: the probability of h (the hypothesis or claim in question) given e (the evidence) – multiply one's estimate of the *inverse* probability, the probability of e given h, by one's estimate of the *prior probability*, the probability of h itself, all divided by one's estimate of the *evidential* probability, the probability of the evidence alone – where one's probability estimates still demarcate the odds one is personally willing to bet on.[17] The posterior probability that it will rain given this type of cloud structure, for instance, is determined by one's probability estimate of this type of cloud structure given that it is raining times one's estimate of the *prior probability* of it raining at all divided by one's probability estimate of this cloud structure obtaining. What results, then, is 'a perfectly objective logic of inductive inference, whose "premises" can be just those prior probabilities, with Bayes's Theorem as the inference engine generating a valid conclusion: the posterior distribution'. Just as in deductive logic, they add, 'you choose the premises, and the engine generates the valid conclusions'.[18]

Note, though, that these 'valid conclusions' obtained by 'objective logic' are limit cases: mathematical ratios of infinite sequences. How then do Howson and Urbach move from *these* to a degree of credence in an item of the sequence without committing a fallacy of division? What justifies the inference from P*(a) to $P(a_o)$? In reply, they provide two arguments.

The first argument is a weaker version of Lewis's deductive justification for the PP: that '(a) objective probabilities [P*(a)] are numerical measures of tendencies scaled in the closed unit interval [0 to 1] and (b) that personal probabilities [$P(a_o)$] reflect personal assessments of relative likelihoods, strongly suggest that conditional

[17] Colin Howson and Peter Urbach, *Scientific Reasoning: The Bayesian Approach*, Third Edition (Chicago: Open Court, 2006), 237. They add here that Bayesians start with a prior probability distribution, which 'reflects a person's belief before the experimental results are known' – 'subjective, ... illusive, idiosyncratic ... likely to vary from person to person' (237). But that said, they then stress that the formula, ideally, and in fact generally, is sensitive to the accumulation of evidence – the observed frequencies or other evidence – and yields an idealized frequency.

[18] Ibid. 265.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

on your sole information that the data source possesses a definite tendency to produce outcomes as measured by P*, your personal probability of an *a*-type outcome [$P(a_o)$] should be equal to P*(*a*)': $P(a_o)|(P^*(a) = r) = r$.[19]

But why 'strongly suggested', why 'appropriate'? For the inference from P*(a), the property of a sequence, to a belief regarding an item in the sequence, Pa_o, is a textbook instance of the fallacy of division. As with Lewis and Ismael, so with Howson and Urbach, there is a magical jump from one's credence in the overall projection, P*(a), to one's credence in any single case within the overall projection, $P(a_o)$. That one's credence is subjective, moreover, does not matter. For while it is certainly the case that one's credence should match what is objective, the objectivity that one's credence should match here is the property we claim to know – the limit-case frequency – and not a component of that property, the single case.

Howson and Urbach do not stop here, however, but offer a second argument, one that makes the much stronger claim that the PP is required by probabilistic reasoning: without it 'the subjective probability formalism would be empty of any methodological content'.[20] For, they ask, how can 'finite sample data … provide information about a chance parameter *even when this parameter refers to an in-principle unobservable limit*'?[21] And, similarly (and prior) to Ismael's strategy of connecting the downward link with the upward link, they reply that it is only through the PP that the 'sample data … [can] provide any information at all about objective probabilities'.[22] How the sample data can provide such information, and where the PP necessarily fits into this process, takes 'three steps' to show – first, 'to recall that … objective probabilities satisfy the axioms of the probability calculus'; and, second, to note that these axioms are framed by '*Kollektivs* [as] Bernoulli sequences [limit cases defined as so]: the probability of any outcome is independent of the outcome at any point, and *is constant from point to point*'[23]; and, as the 'third

[19] Ibid. 76. And later, given P*(a) = r, 'then, in the absence of any other relevant information, the appropriate subjective degree of belief that the event will occur is also r' (174). Whether the 'data source possesses a definite tendency to produce outcomes as measured by P*' will be considered in the discussion on propensity theories.

[20] Ibid. 77.

[21] Ibid. 78 (their emphasis).

[22] Ibid. 77. They add that this capacity of the sample data to provide such objective information at first glance 'seems frankly almost beyond belief'.

[23] Ibid. (my emphasis)

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

and final step', to use 'the resources of [their Bayesian] theory of epistemic probability', the PP, and Bayes's Theorem, to compute the posterior probability. By this they aim to demonstrate that the idealized frequency they derive (and continually improve on) tends to hone in on the objective frequency. Hence, in the end, Howson and Urbach see the PP not as a merely appropriate belief given objective probability claims but as a necessary belief, grounding the very possibility of probability claims.

But, first, there is something strange in the above argument in that it suggests a disbelief in what is being argued for. For note that Howson and Urbach could have seemingly justified the PP by stopping at the second step – the downward link. For this step defines the properties of a Bernoulli sequence, in which the probability of any outcome *is constant from point to point*. If, though, the probability of any outcome were constant from point to point, any application of PP – any inference to $P(a_o)$ from $P^*(a)$ – would immediately follow. For $P^*(a)$ is a Bernoulli sequence that is arrived at by Bayes's Theorem. Hence, if a Bernoulli sequence is the primary conclusion of Bayes's Theorem, and if a Bernoulli sequence is such that the probability of any outcome is constant from point to point, what follows is the PP. But Howson and Urbach rightfully refrain from making this deduction. For Bernoulli sequences are high mathematical abstractions, occupying a platonic mathematical world far removed from our empirical becoming. Hence, to justify the PP by appealing to an attribute of a Bernoulli sequence either would equivocate on probability – to switch from a high mathematical abstraction of probability as defined by a Bernoulli sequence to probability in ordinary life – or would commit a fallacy of division – to infer a property of an item in an idealized empirical sequence from a property of that sequence as a whole.

And so, Howson and Urbach add the third step, the step that brings us into the fluidity of daily reasoning, where they argue that the PP is necessary to reason to a Bernoulli sequence – an 'upward link'. But is this so? To see why not, we might refer back to the point made regarding Ismael's upward link – that one does not need probability to apply to the single case in order to infer the probability of a collective of single cases – the limit case. Bur more, even if we are Bayesians and 'shooting from the hip', even then, and in the spirit of Howson and Urbach, we should aim to frame our subjective estimates in terms of a (Bernoulli) idealized probabilistic sequence – stating our bet on what would happen in the long run – that this single case (a_o) fits into a collective that *as a whole* is defined by $P^*(a)$. That

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

is, even when we estimate the probability 'of' some single case, our estimation is not *of* – does not directly bear on – that case itself but is *of* – directly bears on – the limit-case sequence into which that type of case fits. As von Mises noted, and noted in discussing Bayes, it is 'not too reprehensible' to affirm with '"great certainty" ... that this event will occur on *one* specific trial ... as long as we realize that it is only an abbreviation and that its real meaning is that the event occurs almost always in an infinitely long sequence of observations'; for without such an empirically grounded meaning, probabilistic claims would 'lose their relation to reality' and would be relegated to a mere subjectivism – precisely what Howson and Urbach work so admirably to overcome.[24]

3.

What we have found so far is that the attempts to answer the CA by establishing a legitimate inference to the single case – via a scientific law or a Bernoulli sequence – only make clear just how much of a magical jump such inferences make and how much we want to make – how far we are willing to go to rationalize – such inferences. Another way out, as we have seen, would be to ascribe probability to the single case itself. But how can this be done without succumbing to a mere subjectivism? Propensity theorists answer this question by advancing probability as causally located in the single case.[25] Hence

[24] Richard von Mises *Probability, Statistics, and Truth*, second revised English edition prepared by Hilda Geiringer (London: George Allen and Unwin, Ltd., 1957), 116 and 125

[25] As David W. Miller – in his *Critical Rationalism: A Restatement and Defence* (Peru, IL: Open Court, 1994) – writes: 'Propensities depend on the situation today ... Only in this way do we attain the specificity to resolve the problem of the single case' (186). However, not all propensity theorists hold that propensity is of the single case, whether understood as an object (Peirce), the experimental set up (Popper) or the light cone of one's place in the universe (Miller). For instance, Donald Gillies – in his 'Varieties of Propensity', *British Journal of Philosophy of Science* **51** (December, 2000), 807–835 – argues that propensity does not lie in the single case but in the sequence as a whole. He nonetheless believes that probability applies to the single case by in effect accepting a version of the appeal to appropriateness as proposed by Howson and Urbach in their 1993 second edition of *Scientific Reasoning*, where they argue that 'single case probabilities ... are not themselves objective' but 'subjective probabilities, which considerations of consistency nonetheless dictate must be set equal to the objective

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

there is no jump from probability to the single case for there is an underlying propensity, an objective tendency, *in* the single case that is manifested by some probability – a probability that over the long run comes out as a (dis)confirmable frequency.

That said, one question that immediately comes to the fore is just what propensity is. As Clark Glymour noted, propensity seems to offer nothing more than what is 'revealed in a long-run … sequence of experimental trials of the same kind'. Like Moliere's dormitive virtue in his 'The Imaginary Invalid', which lampoons a group of physicians for purporting to explain why sleep follows the intake of opium – because it has a dormitive property – propensity seems to be merely repetitive, *an empty term*. So, Glymour concludes, propensity is 'a physical property that cannot be recorded and does not necessitate or preclude any occurrence', but merely repeats what the idealized ratio of the aggregate tells us.[26] There seems to be no reality, no 'it', to it.

Recently, Mauricio Suárez addressed this problem by framing his answer 'squarely in Peirce's pragmatist tradition',[27] thereby freeing him from having to justify the idea of propensities by some positivistic reduction of propensities to empirical observations. Rather, they are 'hypothetical … the result of inference' to the best explanation – an explanation that demarcates 'a new and *sui generis* relation' that specifies the 'underlying dispositional properties of things, with

---

probabilities just when all you know about the single case is that it is an instance of the relevant collective' (813). Again: an inference from the property of the collective to that of an item in the collective.

[26] Op.cit., note 16, 284.

[27] Mauricio Suárez, 'Propensities and Pragmatism', *The Journal of Philosophy* **CX** (2013), 68. Suárez grounds his view on five Peircean-inspired pragmatic maxims, in brief: (1) Understand concepts in terms of their effects, but also (2) avoid the positivistic simplicities that identify hypothetical entities with their empirical manifestations; (3) see 'philosophy as continuous with science … bringing its own specific techniques to bear on scientific problems'; (4) 'avoid reducing causal efficacy and causation to anything empirically accessible, such as frequency or correlation'; and (5) 'be open to those scientific theories that postulate hypothetical or fictional entities' via '[a]bduction or … "inference to the best explanation"' (69–72). Suárez also focuses on 'Humphrey's Paradox', where he rejects what he calls the 'identity thesis' that equates probability and propensity – holding instead that 'propensities … are distinct from their probabilistic manifestations' (89, his note 47) and thus need not, like probability, go 'both ways'.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

frequencies as their empirical manifestations'.[28] Such manifestations are 'probabilistic displays' with 'a particular result' – as, say, *this* fair die manifests itself as a probability distribution of 1/6 of a named side coming up.[29] Such probabilistic displays are in the single case: they are expressed 'as probabilities in every single experimental trial'.[30] Hence, Suárez distinguishes '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)'. While the propensity is not observable in the single case, it explains and so grounds the probability of the single case – namely, 'as the exercise of the underlying disposition, which displays itself in a probability' both of which in turn are revealed in the consequences, a long-run frequency distribution.[31]

The question facing any such understanding of propensity is whether it escapes Moliere. To see why not, note that there is an almost exact analogy between Moliere's sleep following the intake of opium/dormitive virtue pair and Suárez's probability/propensity pair. Hence to reject the first pair is also to reject the second.

First, each pair has a side that is an observation and a side that is a purported explanation of the observation where, though, the observation and its purported explanation mutually justify each other in such a way as to merely repeat what the other says. For instance, Suárez holds that atomic decay in the single case 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'.[32] The particles are observed to decay at this rate because they have a propensity, a *half-life*, to do so. And why do we believe they have this half-life? – because they decay at this rate. And why do we believe they decay at this rate? – because they have this half life. But this is no different from saying that we observe

[28] Ibid. 73. Or: They are 'theoretical properties ascribed to objects by scientists in an attempt to explain phenomena involving those objects as dispositional properties with probabilistic displays or manifestations' (87).

[29] Ibid. 87 (Suárez's emphasis). While Suárez initially presents propensities as powers or causes, even mentioning 'Aristotle's efficient causation' (68), he later, on pragmatist grounds, further refines this and asserts that the relation need not be causal. 'It is an empirical matter to determine, for any particular property whether it is causally related to its manifestations' (88). How a manifestation relates to a propensity that does not cause it he does not say, perhaps relying on it being a *sui generis* relation.

[30] Ibid. 76.

[31] Ibid.

[32] Ibid.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

that sleep follows the intake of opium because it has a dormitive virtue. And why do we believe it has a dormitive virtue? – because sleep follows the intake of opium. And why do we believe sleep follows the intake of opium? – because it has a dormitive virtue.

Second, that the observation and its purported explanation merely repeat each other can be further illuminated by the insight that the observed frequency and the disposition purporting to explain it are explained by the same thing – some underlying set of properties. To use Suárez's own example of fragility as the underlying disposition explaining 'easily breakable', why glass is fragile is explained by the same underlying properties as explaining why glass is easily breakable. Why the intake of opium is followed by sleep is explained by the same underlying properties as explain the dormitive virtue of opium. So, too, why some frequency has a certain probability is explained by the same underlying properties that explain why that frequency has a 'propensity' to display such a probability. More, to appeal to an underlying propensity to explain a quantum mechanical probabilistic event is to revert to Einstein's insistence that there must be some cause underlying the event itself – albeit in this case a 'cause' that merely repeats the idealized probabilistic sequence. Hence, any appeal here to Peircean abduction or inference to the best explanation provides no relief. For when the explanation and the to-be-explained are explained by the same thing, then the standards of Peircean abduction themselves reject the explanation as advancing our understanding.

Finally, the analogy between dormitive virtue and propensity is only almost exact in that to say that *propensity* adds an explanation and is manifested in P*(a) = r is worse than saying that *dormitive virtue* explains opium being sleep inducing – or to say that *Energeia* explains and is manifested in $E = mc^2$. While these latter two are each scientific danglers – doing no work, unifying no formula – at least dormitive virtue and *Energeia* call for us to seek out the underlying properties that would explain them. But there is no hunt in the offing for the physical properties of a *sui generis*, whether propensity or probability – no charge, no spin … – nothing to be searched for, as we hunt, say, for the physical properties of a boson or a meson – or as we might hunt for the physical properties of *dormitive virtue* or *Energeia*.

4.

Nonetheless, probabilities, irregular regularities, don't just happen, and that they don't just happen is one of the key reasons backing

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

the conviction that there must be some *one thing*, perhaps even a *sui generis* thing, to explain the regularity that contains the irregularity. As Alan Hájek notes, 'It is not as if facts about the collective impose some constraint on the behavior of the actual sequence … something that *would* impose such a constraint – probabilistic, to be sure – is a single case probability that is fixed from trial to trial. For example, we explain the fact that our coin landed Heads roughly half the time in 1,000 tosses with the hypothesis that its single case probability of Heads is roughly $\frac{1}{2}$, constant across trials, and that the trials are independent of each other'.[33] But if this means, as it does for Suárez, that there is an underlying cause – and the name of that cause is 'propensity' (for Suárez) or 'probability' (for Hájek) – then it only invites the mirth of Moliere.

Why such regularities don't just happen and yet are not to be explained by some underlying propensity (or probability) can be unearthed by harking back to Aristotle's analysis of *spontaneity* and *chance*. Aristotle took spontaneity to be a 'third class of events', designating what comes to pass but 'neither of that which is by necessity and always, nor of that which is for the most part', where chance, he added, is spontaneity within the framework of purposeful action – as a surprise meeting in the market.[34] 'Both belong to the mode of causation "source of change", for either some natural or intelligent agent is always the cause; but in this sort of causation the number of possible causes is infinite … [and] have in fact been caused by something accidentally'.[35] Thus to say that a spontaneous event is 'unaccountable' is correct, for it has a bewildering array of causes, so many and various as to be beyond measure, that merge into a single case – that chance meeting in the market. Aristotle consequently did not

[33] Alan Hájek, "Fifteen Arguments Against Hypothetical Frequentism', *Erkenntnis* **70** (2009), 225. Hájek makes the same point in his predecessor essay – "'Mises Redux' – Redux: Fifteen Arguments Against Finite Frequentism', *Erkenntnis* **45** (1996), 209–227: 'I think the frequentist has things backwards. Surely it is the coin's probability of landing heads that gives rise to its statistics, rather than the other way around' (216). 'Gives rise to' – as if probability is some underlying cause?

[34] Aristotle, *Physics*, trans., R.P. Hardie and R.K. Gaye, in *The Complete Works of Aristotle: The Revised Oxford Translation*, ed., Jonathan Barnes, Volume One, (Princeton, New Jersey: Princeton University Press, 1984), 196:12–13. Aristotle was not thinking of probability, of which he had not a glimmer, when he spoke of what comes to pass for the most part, but of contingency – as in a boulder not falling because it was lodged on the edge of a cliff.

[35] Ibid. 198a 2–3

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

see spontaneity as a 'unique and *sui generis*' cause in itself but rather took it as a term of art 'about' causes, like the term 'means of transport' is 'about' an indefinite number of ways of transportation – hobby horses, scooters, carts, bikes, cars, trains, planes, horses, strong backs, … – with no clear closure but designating in the case of 'spontaneity' that vast array of causes of events that are individually explicable up to a point but beyond any general accountability.

Propensity, I submit, is like spontaneity, 'belonging to the mode of causation "source of change"', but now a *fourth class of event*. While 'spontaneity' designates an array of multiple and varying causes, this or that or whatnots, that unpredictably come together to bring about a single case, an irregular irregularity, 'propensity' designates an array of multiple and varying causes that unpredictably come together to bring about a single case, but over time bring together this type of single case in a way that hovers around some idealized ratio – an irregular regularity.[36] Neither concept – spontaneity or propensity – scientifically explains anything, for neither designates any specific underlying cause or primitive that we can further investigate. Both are abstractions that designate a hodgepodge of an indefinite and varying number of causes.[37]

Why, for instance, did the die come up six on this throw? On the one hand, we could say with Aristotle that the number of causes is *indefinite*, and so 'unaccountable' – the dirt on the die, the wrist movement of the thrower, where and how the die was gripped, how the thrower followed through on the throw, the surface the die landed on, and where, and at what angle, the density of the air itself, … perhaps even, to jump from Aristotle to Epicurus, the 'swerve' of an uncaused event. On the other hand, we could say with Hume that the event is uncaused, that 'chance is nothing real

[36] Whether it is possible to put the happenings of irregular irregularities into an idealized sequence I will leave to those who have the courage to define 'this or that or whatnots' – or even 'market surprises'.

[37] This is not to deny that probabilistic laws provide a scientific 'explanation', just that propensities as underlying causes of such laws provide such an explanation. Just how probabilistic laws explain, of course, is another issue. But even if, as Wesley C. Salmon argues – in his 'Explaining Things Probabilistically', *The Monist* **84** (2001), 208–217 – probabilistic laws explain the single case, they are not effective in the single case. As Salmon notes, 'statistical explanations – in many cases, at least – are *not* arguments' (210). For explanation goes one way, from the single case to the law; it is after the fact. In contrast, a guide, like a prediction or argument, goes the other way, from the law to the single case; it is before the fact – a measure we do not have.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

in itself, and, properly speaking, is merely the negation of a cause, its influence on the mind is contrary to that of causation'.[38] For when we find no regularity other than the framework of the irregular regularity itself, we have no specific causal, i.e., predictive or accountable, factor to appeal to in the single case. What we have instead is an indefinite but finite, relatively closed, interconnected, and exceedingly complex system of variously weighted causes, such that, no surprise, there will be these recurring long-run idealizations just because there are only so many options for the world causes to work out as they do.[39]

In sum, Hájek is quite right to insist that what we want to know is the probability of the single case, one's own case: 'the *fact* that [a person] ultimately cares about is a fact about himself – one expressed by a meaningful "probability of death" statement that refers to a single person'.[40] But this desire is ambiguous and is in danger of blurring the problem of the reference class with that of the single case. A fact that we legitimately desire to know about ourselves is what precise reference class – what particular distribution of possibilities – we fit into. But even knowing that does not tell us what we really, and, alas, illegitimately, want to know – namely, what our 'single-

[38] David Hume, *Treatise on Human Nature,* ed., L.A. Selby-Bigge (Oxford: Clarendon Press, 1967), 125.

[39] This understanding of propensity as involving multiple and varying causes, admittedly, will not address quantum mechanics. Henry E. Kyburg – in his 'Propensities and Probabilities', *The British Journal of the Philosophy of Science* **25** (1974) – notes the striking difference between radioactive decay and, say, dying. 'If we could identify people chemically and physically so as to put them in a small number of classes, and that is *all* we could do, then we would be home free' in taking radioactive decay and dying as similar. 'But precisely in distinction to the decay of radioactive elements, we can often *identify* the cause of a person's demise' (366). The point is, if we could identify the cause of the half-life, as in discovering the underlying properties of a 'propensity', we would have left the swerve of quantum mechanics for some underlying Einsteinean cause. And, whatever, probability would not be of the single case except elliptically.

[40] Alan Hájek, op. cit, note 33, 218. Similarly, Christopher Hitchcock – in 'Causal Generalizations and Good Advice,' *The Monist* **84** (2001), 218–241 – starts with the PP as a 'useful framework for thinking about the relationship between subjective probabilities and beliefs about objective probabilities' (235) but adds that the PP needs a supplement, an answer to the reference class problem, by insisting that the probabilistic law is framed as directly as possible to the specific case so that there is less of a gap between them. But a gap in this instance is a chasm – not to be bridged by a magical jump from the real probability of the aggregate to the imaginary probability of the single case.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

case probability' is – as if probability is a cause, or has propensity as a cause, that works to favor us in our single case. Even with getting the reference class exactly right, in other words, the single case has no probability of its own but remains, as randomness requires, a happenstance within the probability of the sequence as a whole. With Aristotle, as with Hume, a concept like chance demarcates the limits of explanation and is no explanation itself. While this is not to go so far as Glymour, who relegates probability to the netherworld of 'Locke's substance … something we know not what',[41] it does limit our understanding of probability to a class of varying and myriad causes coming together in a relatively closed and finite universe.

5.

If probability does not apply to the single case, how, if at all, can it be a guide in our lives? This question expresses what I take to be the deepest motivation for single-case probability reasoning: probability really is a guide in our lives and as a guide is so in the single case. How, then, do we bring together these three necessities – that probability is a guide in our lives, that we act in the single case, and that probability does not apply to the single case? The answer, *no matter what view of probability we take* – law-based, Bayesian, even propensity – is to distinguish between being a guide in the single case and applying to the single case.

First, probability guides us in the single case by framing single cases within (dis)confirmable idealizations of irregular regularities – where frequentists insist that their idealizations are to be tested, after Bas C. Van Fraasen, by 'vindications' like the *Brier Score*: a combination of 'informativeness or extremeness' – the closer to 1 or 0 the higher the score; and by 'calibration' – 'that the forecasts fit the series of actual events perfectly', e.g., that it really does rain 40% of the time on a 40% forecast;[42] where Bayesians, like Howson and Urbach, stress that their instantiated Bayesian formulas are to be improved by subsequent observations or reasons, always placing probability within a Bernoulli-defined collective; and where

[41] Op. cit., note 16, 299.

[42] Bas C. Van Fraasen, 'Calibration: A Frequency Justification for Personal Probability', in *Physics, Philosophy and Psychoanalysis: Essays in Honor of Adolf Grünbaum*, eds, R.S. Cohen and L. Lauden (Boston: D. Reidel Publishing Company, 1983) Boston Studies in the Philosophy of Science, vol. **76**, 300.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

propensity theorists, like Suárez, insist that any propensity be 'displayed' and so tested by its long-run consequences.

That is, probability guides us in the single case by underlining that, however we interpret probability, we must ground it in some type of long-run projection – from idealized observed frequencies to Bernoulli sequences to propensity-driven observed consequences. The abstract mathematical coin flip by itself is no practical interest to us other than to model what we discover what actually happens in real coin flips. Whatever we might think of Molinieux's blind person seeing a cube for the first time, until we see a sequence play out over time, the mathematics doesn't transfer into our experience. First came the games of chance, the empirical die rolls, and only after that the mathematical models, and only then, with the owl of Minerva, the interpretations of probability. And what these models and interpretations give us, as I have argued, and as professional gamblers insist, is not anything in regard to the single case but projections – namely, what is to be anticipated in the long run, whether in terms of laws, Bayesian formulas, or propensity-driven consequences.

Hence, and second, probability guides us *negatively* by *warning* that chance is the inapplicability of limit-case sequences to the single case. Chance serves as a warning of the vagaries of irregular regularities – that there is no or no known algorithm within the sequence. To see chance as an underlying reality gives us the false confidence that the odds are in our favor *here and now*, when all the odds do – to repeat, as lawful or Bayesian or propensity-driven (dis)-confirmable idealizations – is cover the aggregate.

As a result, and paradoxically, probability guides us by underlining that the one crucial idea behind probability is to avoid applying probabilistic credence to the single case at all. Contra Lewis, the grounding principle of probability is not the *Principal Principle* but the *Anti-Principal Principle* (APP): chance does not apply to the single case and thus our probabilistically-based decisions, our 'bets', should never be focused on the single case. Professional gamblers see this with eyes wide open. As Orange County KO, a former black jack professional gambler, put it, 'Serious players ... measure their result in terms of EV [the normative expected value of the game] ... a simple formula: EV = overall edge * average bet * hands per hour * hours played ... [where, following this] you can rest assured that your losing streak is simply a result of negative variance ... let the math take care of itself'.[43]

[43] Orange County KO, 'Mastering Your Reaction to Losses', *Blackjack Forum* ***XXIII*** (Spring, 2003). While this is expressed in terms of a

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

There is a profound difference between the PP, allowing us to think that one has a good probability *in this case* – the die's coming up not-six – and the APP, or better still, its amplified version, the EV, that the die will come up not-six about five-sixths of the time over the long run. This is a conceptual as well as a psychological point: to think in terms of the PP is to think in terms of the single case rather than a long-run projection. It is to put the weight of one's bet, the reliance of one's judgment, on something that does not exist, namely, that probability is working in one's favor here and now – say, *the good five-sixths chance* that six won't come up. To think in terms of the APP and EV is to be made acutely aware of the vagaries of a five-sixths ratio and the necessity of being able and willing to stay for the long run. It might be more fun and exciting to view chance as operative in the single case. But it is no way to manage one's life. One should play the long run or not at all. As another professional gambler, 'Jimmy', put it, 'Amateurs expect to win – after all, you're supposed to win when you have an edge, right? ... This is a monumental mistake ... The point is this: anything is possible on a given day – after all, it's gambling. ... Some days you'll lose, some days you'll win – it really doesn't matter because in the long run you'll always win'.[44]

6.

Jimmy's bald claim that in the long run you'll always win, of course, comes with many caveats – that one must have the frequencies right, that one can play long enough to endure a bad run and thereby take advantage of the long-run ratios. Yet in addition to these caveats, there are two other and much neglected caveats to face and, as well, to appreciate how probability guides us in our lives. The first is the *one-off problem*: most of the important decisions we make in life are one offs, events that are not just single cases of the same type on which we can project but single cases that are seldom if ever repeated in our lives. How, then, can we stay the course on them?

Take, say, the hard case of deciding on a surgery offering a substantial improvement in one's life but with severe consequences if the

---

frequentist understanding of probability, it could equally well be expressed in Bayesian or propensity terms.

[44] As interviewed by Justin Rohrlich, in 'Hunches Mean Squat: How a Professional Gambler Invests in Sports', *Minyanville* (July 10, 2012).

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

surgery goes south. The surgeon reports that the operation has a 99.9% success rate – one out of a thousand goes wrong – and confidently assures us of our 'good chances'. How though are we to understand this? What we dearly want to say is precisely what we should not. What we dearly want to say, and so engage in magical thinking, is that we, individually, have a 'good chance' of coming out ahead. As if there is *something* – a law, a Bernoulli sequence, a propensity, the gods of chance – working on our own behalf in this case, *one's own* case. But, after the CA and the Anti-PP, that's not so. What we then often do is try for the second best option, *the surgeon's point of view*, the view of the one who can say that of the patients s/he treats by this operation, 99.9% come out well. But the surgeon's point of view is not available here either, just because f*or each of us* there is no projection. For each of us it's a one off, seldom if ever, to be seen again. In this case, as in so many others, *we* individuals are the single case itself.

When I made this observation to my surgeon she immediately understood. But then I noted that she would regard me as a nutter if I did not accept her probabilistic advice on the procedure I faced. And she agreed with that too – seeing, I believe uncomfortably, the apparent contradiction. How to resolve it, and, I suggest, how we in practice implicitly do resolve it, is to take the *ordinary agent's point of view* and *bundle* together items with similar long-run projections. The principle of bundling is this: take the one offs with similar projections and bundle them together with all such projections. This in effect gives us a surrogate of the surgeon's point of view.[45]

[45] As reported in *Boston Globe* (2/2/14, A17), the risk of death for donors of part of their liver is one to five for every 1,000 transplants. Dr. Giuliano Testa, Baylor University's Dallas Medical Center's surgical director, is cited as worrying that too much focus on rare donor deaths will decrease the number of donors and so cost lives. In a classic instance of the surgeon's point of view, he observed, 'Bad things happen all the time. They should not take center stage with the discussions we have with the donors'. No matter how careful you are with regulations and donors, 'you still cannot totally control unforeseen circumstances'. Clearly, this is the read of the one who is in it for the long run, who knows that most of his bets of this type by far will come out right. (Does it also provide a backdrop for why surgeons are notoriously unsympathetic? – that, like Jimmy, they know it's gambling, after all?) The patients, in sharp contrast, are well advised to know that the long odds of the surgeon are not for them. This is not to say that one should not donate, but let's not downplay what the bet is on, a *bundled frequency*, and the profound altruism in such donations. (My thanks to Charlene Entwistle for this reference.)

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

Similar projections allow us to bundle different one offs into an overall projection and so stay the course for them. The mathematics is the same and the mathematics is all that counts, provisos in place. Bundling allows us to have long-run projections of vastly different and unrepeated types of situations that more or less share the same (dis)confirmable idealized frequencies. In this way, then, we can play the odds over the long run even though for most of the time we are addressing one offs. There are three central ideas back-boning the practice of bundling.

The first idea is that we must be able to discern the type in the token event, the one off, and thereby legitimately estimate how it would play out in time. Take the problem of the newly minted coin – tossed once and then destroyed. What is its probability of landing heads? My answer builds on Kyburg's reasoning that we should favor our observation of how all coins but quarters land over our observation of how quarters land – namely that coins, being more numerous than quarters, offer a greater evidential base, and so, given that they, coins and quarters, are sufficiently alike to land similarly, we should take it that the data we have for all coins provides us with a better understanding of how quarters will land than the data on quarters alone. There is no reason, he says, to take quarters as 'weird' and so we should see them as landing heads in the same ratio as coins in general do.[46] By parallel reasoning, Kyburg continues, as with quarters so with the newly minted coin: we have no reason to take it as weird and so should regard it as of a type with all coins. What Kyburg is in effect doing here is bundling the newly minted coin with all coins, and, to extrapolate further, with all such projections with a limit of $\frac{1}{2}$. In this way, we can place or bundle the one off, the newly minted coin, in a long-run sequence, however understood.

The second idea is that bundling pays off in the long run – even if all we project on are one offs. Take, for instance, that famous (or once-famous) one off found in *Let's Make a Deal,* where the host, Monty Hall, at the end of the game asked the contestants to choose one of three doors, one of which had a significant prize behind it. After they chose a door, he then opened a door with no prize behind it and then offered them the option of switching their choice to the door they did not choose. Should they do so? If they were playing

[46] Henry Kyburg, 'Don't Take Unnecessary Chances!' *Synthese* **132** (2002), 11. This also provides the basis of a reply to speaking of the probability of the universe coming to be – by not seeing it as unique but as a type. While Kyburg speaks here in frequentist terms, I could just as well have spoken in Bayesian or propensity terminology.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

this game over the long run, if they were the surgeons in this game, they should, surprisingly to most of us, switch; for there is a $\frac{2}{3}$ probability that *over the course of time* they would win twice as often as they lost.[47] But the contestants have no sequence to play out here. So, should they switch? They should, but not because the odds favor them in this single case. Since they will never face this choice again, there are no odds in this case as an isolated instance just because there is no projection – no law, no Bernoulli sequence, no propensity – for them to think in terms of the long run. But by bundling this case with all the other $\frac{2}{3}$ projections they will face in their lives – whether one offs or repeated, whether hands of poker, bridge, gin rummy, insurance plans, partners – they and we can see that they will indeed win more than they lose. Not to bundle is to be the nutter the surgeon would have taken me to be had I not heeded her advice, for not to bundle, to treat each one off as an isolated single case, is to doom oneself to coming out the short end more often than not in all the one offs one faces. The bundled long-run projections – as laws, Bernoulli sequences, propensity-driven consequences – rule.

7.

The third idea behind the practice of bundling is to rein in any false optimism the practice might engender. For, back to the operation, and now given the practice of bundling, what concentrates our mind here is that what works in this one off, what we are projecting on, is, not *this* operation, but all those one offs with 99.9% projections that we will face throughout our lives, and that, while acting on those

[47] If they had been offered the choice of picking two doors or one, they would have picked two; and the odds would not have changed on those two doors winning two thirds of the time even after one of them had been opened. See Peter Baumann, 'Three Doors, Two Players, and Single-Case Probabilities', *The American Philosophical Quarterly* **42** (2005), 71–79. Bauman sees no way out here, arguing that 'one ought to give up probabilistic arguments for or against switching or sticking in isolated individual cases. They only make sense in the long run' (76) – which I disagree with because of the practice of bundling. For a delightful discussion of the authoritative vitriol directed against the advice to switch as offered by Marilyn vos Savant of *Ask Marilyn* fame, see Leonard Mlodinow, *The Drunkard's Walk: How Randomness Rules Our Lives* (New York: Pantheon Books, 2008), 42–45.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

projections assures us that we will 'win' far far more than we 'lose', we have no idea which of these situations will leave us with a loss – a life, a limb, a house, a donut. Whether we will lose the life or the donut is unknown – other than that we will lose some things or other things over the course of time with that projection.

That is, what concentrates the mind here is that one offs like surgeries are *forced bad-odds one offs* – choices that are not elective but unavoidable and with odds that we would almost never accept in a game. Such choices starkly illuminate the truth behind Peirce's warning 'that every gambler, if he continues long enough, must ultimately be ruined', on the grounds that eventually one would be hit by a deviant run that one could not overcome.[48] But, even more sobering than Peirce's worry of a bad streak, *even always winning without enduring any bad streak*, we still face the risk of forced bad-odds one offs. We don't project just on games but also on our health, family, spouse, partner, friends, colleagues, financial advisor, the market, the national arena, the international arena ... In such situations we can't protect ourselves by choosing which 'game to play' just because these are not games but are types of cases – *live* and *forced* after William James – that we must project on if we are to live full lives. Eventually then, and generally in a much shorter time for most of us than Peirce's deviant run, there would come a series of rational choices on forced bad-odds one offs that turn out so badly that one could not or would be hard put to recover.[49] And once we realize this, we see into Jimmy's *long-run winner's fallacy* – that, provisos in place, 'in the long run you'll always win'. You will, but not necessarily in the games you 'must' win. Those of us who have been 'fortunate' have indeed, as all

[48] Charles Sanders Peirce, *Philosophical Writings of Peirce,* ed., Justus Buchler (New York: Dover Publications, Inc., 1955, 169. In his words, a gambler – which here includes everyone – 'cannot be absolutely *certain* that the mean result will accord with the probabilities at all. Taking all his risks collectively, then, it cannot be certain that they will not fail,' for that is what *random* means (ibid).

[49] Ibid. 162. It is to the greatness of Peirce that he sees the falsehood of the long-run winner's fallacy and offers a solution – that of epistemic community. 'I can see but one solution ... that logicality inexorably requires that our interests shall *not* be limited. They must not stop at our own fate, but must embrace the whole community. ... To be logical men should not be "selfish"' (162). This begins to open up the issue of the morality embedded in reasoning probabilistically and, more generally still, the morality embedded in epistemic vulnerability – of which this piece is a part.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press

should, relied on the projections of these forced bad-odds one offs to come out right, and we have been for the most part *lucky*.[50]

EDWARD JAMES (ejames@bridgew.edu) *is an Emeritus Professor of Philosophy at Bridgewater State University. He has published in such journals as* Philosophy, Mind, *and* Ethics, *with his more recent essay, 'Too Soon to Say' in* Philosophy ***87*** *(2012), 421–442. His work centers on the extent and limits of reasoning as they frame our epistemic vulnerabilities – where our best laid plans so easily go awry or are met by sharp critique.*

[50] My thanks to my philosophy department, at Bridgewater State University, where I read a much earlier version of this at a Philosophy Colloquium, and was helped, as always, a great deal. Thanks especially to Professors Matthew Dasti, William Devlin, Robert Fitzgibbons, Gull Kober, Laura McAlinden, James Pearson, Aeon Skoble, Catherine Womack – and, as always, especially, to Professors David Cheney and Steven Sanders. The errors, of course, are mine.

https://doi.org/10.1017/S0031819115000121 Published online by Cambridge University Press