

HAVE YOUR CAKE AND EAT IT TOO:
THE OLD PRINCIPAL PRINCIPLE RECONCILED WITH THE NEW

Peter B. M. Vranas
vranas@iastate.edu
Iowa State University
10 July 2002

Abstract. David Lewis (1980) proposed the Principal Principle (PP) and a “reformulation” which later on he called ‘OP’ (Old Principle). Reacting to his belief that these principles run into trouble, Lewis (1994) concluded that they should be replaced with the New Principle (NP). This conclusion left Lewis uneasy, because he thought that an inverse form of NP is “quite messy”, whereas an inverse form of OP, namely the simple and intuitive PP, is “the key to our concept of chance”. I argue that, even if OP should be discarded, PP need not be. Moreover, far from being messy, an inverse form of NP is a simple and intuitive Conditional Principle (CP). Finally, both PP and CP are special cases of a General Principle (GP); it follows that so are PP and NP, which are thus compatible rather than competing.

*1. Introduction.*¹

You have a crystal ball. Unfortunately, it’s defective. Rather than predicting the future, it gives you the *chances* of future events. Is it then of any use? It certainly seems so. You may not know for sure whether the stock market will crash next week; but if you know for sure that it has an 80% chance of crashing, then you should be 80% confident that it will—and you should plan accordingly. More generally, given that the chance of a proposition *A* is $x\%$, your conditional credence in *A* should be $x\%$. This is a *chance-credence principle*: a principle relating chance (objective probability) with credence (subjective probability, degree of belief). Let’s call it the *Minimal Principle* (MP).

Your crystal ball sometimes malfunctions. Not that it makes mistakes; oh no, this never happens. It’s rather that sometimes the ball doesn’t *just* give you the chance of the specific future event you are interested in; it gives you also the chances of *other* future events, and those of some *past* events as well. If “what’s past is no longer chancy”, then the latter chances are always either zero or one; so providing them amounts to recounting the past. No matter: you should just disregard the extra output. It constitutes *admissible* evidence: it doesn’t affect the credence you should assign to your event of interest, which remains equal to what the crystal ball tells you the chance of that event is. More generally, given that the chance of a proposition *A* is $x\%$, and given any admissible evidence (including information about other chances and about the past), your conditional credence in *A* should be $x\%$. Following Lewis (1980), let’s call this second chance-credence principle the *Principal Principle* (PP).

Occasionally your crystal ball goes berserk. Rather than just obediently answering your specific question, it spits out the chances of *all* past and future events. No great harm done: you should again disregard the extra output. More generally, given the chances of all propositions, your conditional credence in a proposition *A* should be equal to the chance of *A*. This is just a consequence of PP. Following Lewis (1994), let’s call this special case of PP the *Old Principle* (OP). Let’s also say that PP is an *inverse form* of OP.²

I am very grateful to Frank Arntzenius, Michael Bishop, Alan Hájek, Carl Hoefer, and Michael Strevens for helpful comments. Thanks also to Paul Bartha, Jack Davidson, Kevin de Laplante, Adam Elga, Heimir Geirsson, Aviv Hoffmann, and Tony Smith for help or interesting questions. This paper was presented at the 2002 Rutgers Epistemology Conference.

¹ This section presents the issues in a simplified and slightly imprecise way. Rigor is introduced in later sections.

² More generally, say that (a chance-credence principle) P_1 is an *inverse form* of P_2 exactly if P_1 entails P_2 and the standard formulation of P_2 , unlike that of P_1 , gives a credence conditional on the chances of all propositions.

Lewis (1994) believes that PP runs into trouble. I will explain his reason later on (§4); for the moment let's focus on his reaction to this belief. He argues that information about present chances always reveals information about the future and is thus never admissible. He concludes that PP "never applies", except as an approximation to the following *New Principle* (NP): your conditional credence in a proposition A , given the chances of all propositions, should be equal to the *conditional* chance of A , given the chances of all propositions (details in §3). This principle, however, is normally useless: to apply it you need the chances of *all* propositions, which you have only if your crystal ball goes berserk. NP is analogous to OP; what you want instead is an *inverse* form of NP, something analogous to PP. Lewis tries, unsuccessfully, to simplify such an inverse form. He thinks it's "quite messy", so that NP is less intuitive than PP. He concludes that a feature of reality deserves the name of chance to the extent it obeys PP rather than NP, "so nothing perfectly deserves the name". Pretty depressing.

My first main thesis in this paper is that an inverse form of NP, far from being messy, is surprisingly simple. It's the following *Conditional Principle* (CP): for any propositions A and B , given both B and that the conditional chance of A given B is $x\%$, your conditional credence in A should be $x\%$. Indeed: if you know that B is true and your crystal ball tells you that the chance of A given B is 80%, then you should be 80% confident that A is true. I will argue that CP entails NP. Is this vindication of NP the *coup de grâce* for PP? Not at all. My second main thesis is that both PP and CP, and thus both PP and NP, are special cases of the following *General Principle* (GP): given that the conditional chance of A given B is $x\%$, and given also both B and any admissible evidence, your conditional credence in A should be $x\%$. PP and NP are thus compatible rather than competing. But what about Lewis's claim that PP never applies because information about present chances is never admissible? My third main thesis is that Lewis's reaction to the putative trouble with PP is an overreaction: even if in certain exotic cases PP doesn't apply, in ordinary cases it applies exactly—not just as an approximation.

In §2 I formulate more formally MP, CP, PP, and GP. In §3 I argue that, just as PP entails OP, CP entails NP. In §4 I argue that PP is not vacuous. In §5 I offer some remarks on admissibility. I conclude in §6.

2. A hierarchy of chance-credence principles.

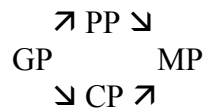
Let Cr be the credence function of a given person (at time t); i.e., the function that assigns to any proposition A about which the person has a belief a number in $[0, 1]$ which is the person's degree of belief in A (assuming, to simplify, that every belief has a degree). Let Ch be the chance function (at t); i.e., the function that assigns to any proposition A which has a chance a number in $[0, 1]$ which is the chance of A . Assume that Cr should be, and that Ch is, a probability measure. Let $\langle Ch(A)=x \rangle$ be the proposition that $Ch(A)$ is x , x being a number in $[0, 1]$. Let B and E be any propositions. Then four of the chance-credence principles which I formulated informally in §1 can be formulated more formally (though still very roughly: see §5) as follows.

- Minimal Principle* (MP): $Cr(A|\langle Ch(A)=x \rangle)$ should be x .
- Conditional Principle* (CP): $Cr(A|B\langle Ch(A|B)=x \rangle)$ should be x .
- Principal Principle* (PP): If E is admissible with respect to $\langle Ch(A)=x \rangle$, then $Cr(A|E\langle Ch(A)=x \rangle)$ should be x .
- General Principle* (GP): If E is admissible with respect to $B\langle Ch(A|B)=x \rangle$, then $Cr(A|EB\langle Ch(A|B)=x \rangle)$ should be x .

Take, for example, A to be the proposition that a given (indeterministic) coin will come up heads at the next toss. (1) If you know (for sure) that $Ch(A)$ is 50%, then according to MP your credence in A

should be 50%.³ (2) CP is one generalization of MP. If, rather than knowing that $Ch(A)$ is 50%, you know both that (B) the coin will be tossed next by a magician and that $Ch(A|B)$ is 50%, then according to CP your credence in A should (again) be 50%.⁴ (3) PP is another generalization of MP. If, in addition to knowing that that $Ch(A)$ is 50%, you know that (E) the coin came up heads at the last toss, then it seems that your credence in A should again be 50%. The information in E that is relevant to A is “superseded” by the information that $Ch(A)$ is 50%; in other words, E is *admissible* with respect to $\langle Ch(A)=50\% \rangle$. This, however, does not follow from PP, which is silent on the question of *which* propositions are admissible. So PP needs to be supplemented with a substantive account of admissibility.⁵ I will say more about this issue later on (§5); for the moment let’s proceed with the intuitions about admissibility that we do have in particular cases (as illustrated by the example I just gave). (4) Finally, GP combines the above two generalizations of MP. If, in addition to knowing both that (B) the coin will be tossed next by a magician and that $Ch(A|B)$ is 50%, you know that (E) the coin came up heads at the last toss, then it seems that your credence in A should again be 50%. GP, like PP, needs to be supplemented with a substantive account of admissibility.

GP entails both CP and PP. To see this, suppose you should be certain that a proposition I is true and has chance one; e.g., let I be the proposition that either 971 is a prime number or it is not. Then replacing in GP E with I yields CP (assuming that I is admissible), whereas replacing in GP B with I yields PP.⁶ Similarly, PP entails MP, and so does CP. We have thus a “hierarchy” of chance-credence principles which can be represented in the following diagram. (I leave open the question of whether further arrows can be added to the diagram.)



3. A partial vindication of the New Principle.

Let ‘ f ’ denote rigidly (Kripke 1980: 77-8) a probability measure. Let T_f be the proposition that Ch is f , i.e., that (i) $\text{dom } Ch = \text{dom } f$ and (ii) for every A in $\text{dom } f$, $Ch(A) = f(A)$. T_f entails propositions (like $\langle Ch(A)=f(A) \rangle$) that specify the chances of various propositions. For example: if $f(A)$ is 0.3, then T_f entails the proposition that $Ch(A)$ is 0.3 (so T_f is false if $Ch(A)$ is not 0.3). Let $f(A|B)$ be $f(AB)/f(B)$. Let H be a proposition that completely specifies the past and present (except for the present chances).⁷ Then the remaining two chance-credence principles which I formulated informally in §1 can be formulated more formally (though still very roughly: see §5) as follows.

³ MP corresponds to “Miller’s principle”. When applying MP, care is needed to avoid “Miller’s paradox” (Miller 1966a; Popper 1966a, 1966b; Mackie 1966; Miller 1966b; Bub & Radner 1968; Miller 1968; Rozeboom 1968; Jeffrey 1970; Good 1970; Howson & Oddie 1979; Howson & Urbach 1993: 398-401; cf. Lewis 1980: 89; Strevens 1999: 270 n. 2).

⁴ CP is formulated by van Fraassen (1989: 201-3; cf. 1980: 106-7; Skyrms 1988). Domotor (1981: 32) and Strevens (1995: 555) formulate similar principles.

⁵ PP is just a consequence of the following definition of admissibility: E is admissible with respect to $\langle Ch(A)=x \rangle$ exactly if $Cr(A|E\langle Ch(A)=x \rangle)$ should be x . This definition does not provide a substantive account because it does not say for *which* propositions $Cr(A|E\langle Ch(A)=x \rangle)$ should be x . (To define *inadmissibility*, replace “should be x ” with “should differ from x ”. If $Cr(A|E\langle Ch(A)=x \rangle)$ may be x but may also differ from x , then E is *neither inadmissible nor admissible* with respect to $\langle Ch(A)=x \rangle$.)

⁶ *Proof sketch.* $\langle Ch(I)=1 \rangle$ entails $\langle Ch(A)=Ch(A|I) \rangle$, so $E\langle Ch(A)=x \rangle \Leftrightarrow E\langle Ch(A|I)=x \rangle$. So if $Cr(\langle Ch(I)=1 \rangle) = 1$, then $Cr(E\langle Ch(A)=x \rangle \Leftrightarrow E\langle Ch(A|I)=x \rangle) = 1$, and thus $Cr(E\langle Ch(A)=x \rangle) = Cr(E\langle Ch(A|I)=x \rangle)$; similarly, $Cr(AE\langle Ch(A)=x \rangle) = Cr(AE\langle Ch(A|I)=x \rangle)$. Then $Cr(A|E\langle Ch(A)=x \rangle) = Cr(A|E\langle Ch(A|I)=x \rangle) = Cr(A|EI\langle Ch(A|I)=x \rangle)$ (given $Cr(I) = 1$).

⁷ More precisely: H specifies, for all time instants up to and including t , the spatiotemporal arrangement of what Lewis (1994: 474) calls “local qualities” (“perfectly natural intrinsic properties of points, or of point-sized occupants of

Old Principle (OP): If HT_f is admissible with respect to $\langle Ch(A)=f(A) \rangle$, then $Cr(A|HT_f)$ should be $f(A)$.
New Principle (NP): $Cr(A|HT_f)$ should be $f(A|HT_f)$.⁸

PP entails OP. Here is why. Suppose that HT_f is admissible with respect to $\langle Ch(A)=f(A) \rangle$. Replace in PP E with HT_f and x with $f(A)$. Then $Cr(A|HT_f \langle Ch(A)=f(A) \rangle)$ should be $f(A)$. But T_f entails $\langle Ch(A)=f(A) \rangle$. It follows that $Cr(A|HT_f)$ should be $f(A)$.

My first main thesis in this paper is that CP entails NP. Here is why. Replace in CP B with HT_f and x with $f(A|HT_f)$. (This assumes that HT_f is in $\text{dom } f$, but so does the very formulation of NP.) Then $Cr(A|HT_f \langle Ch(A|HT_f)=f(A|HT_f) \rangle)$ should be $f(A|HT_f)$. But T_f entails both $\langle Ch(AHT_f)=f(AHT_f) \rangle$ and $\langle Ch(HT_f)=f(HT_f) \rangle$; so T_f entails $\langle Ch(A|HT_f)=f(A|HT_f) \rangle$.⁹ It follows that $Cr(A|HT_f)$ should be $f(A|HT_f)$.

The derivation of OP from PP is well known (e.g., Lewis 1980: 97), but as far as I am aware the parallel derivation of NP from CP is novel (though cf. Strevens 1995: 557).¹⁰ The latter derivation is important because it allays two worries about NP which proponents of NP take quite seriously. First, the worry that NP is less intuitive than PP because an inverse form of NP “gets quite messy” (Lewis 1994: 489). Second, the worry that NP “seems user-hostile, since it requires us to evaluate such esoteric quantities as the *conditional* chance that the coin lands heads, given some complete law-proposition” (Hall 1994: 512). Lewis responds to the first worry by biting the bullet. He claims that chance can be defined as that feature of reality which obeys PP (rather than NP); because PP never applies, “nothing perfectly deserves the name” of chance, but “an imperfect candidate may deserve the name quite well enough” (1994: 489). Both Lewis and Hall respond to the second worry by arguing that in ordinary cases we may go on using PP: it provides a good approximation to NP. My derivation of NP from CP provides better responses to these worries: an inverse form of NP, namely CP, is about as intuitive as PP and is not “user-hostile”.¹¹ So the derivation provides a vindication of NP.¹² (This vindication is only partial, however, because NP relies on the arguably dubious assumption that HT_f is in $\text{dom } f$: is HT_f the kind of proposition that has a chance?¹³)

points”). So although H specifies the “present”, it does not specify the present chances if chance-related properties are not local qualities (cf. Ismael 1996: 80-1). Like T_f , H need not be true.

⁸ My formulation of NP differs slightly from those of Lewis (1994: 487) and Hall (1994: 511): they assume that H and T_f are true (at some possible world) and that $f(H)$ is one (because what’s past is no longer chancy), so $f(A|HT_f) = f(A|T_f)$. (Moreover, their T and T_w differ slightly from my T_f . See also footnote 14.)

⁹ Here I use the “ratio analysis” of conditional chance: $Ch(A|B) = Ch(AB)/Ch(B)$. This analysis might be flawed (Hájek 2002), but I am not committed to it. If, e.g., conditional chances are primitive, then redefine T_f as the proposition that $Ch(\cdot|\cdot)$ is $f(\cdot)$, and my proof goes through.

¹⁰ Strevens derives NP from a principle which in my notation is: $Cr(A|BT_fK)$ should be $f(A|B)$. This principle, although somewhat similar to CP, is *not an inverse form* of NP: it gives a credence conditional on the chances of *all* propositions (that have a chance according to a “probabilistic theory”). So Strevens’s derivation does not quell the two worries about NP that I go on to mention in the text.

¹¹ The inverse form of NP which Lewis is unable to simplify (cf. §4) differs from the inverse form of NP which is CP; but I take it that the existence of even one intuitive and user-friendly inverse form of NP suffices to allay the two worries.

¹² But doesn’t the derivation make NP redundant, given that CP is much easier to use? Not quite: NP is better suited than CP for dealing with certain “undermining futures” (§4).

¹³ Cf. Hoefer 1997: 328; Spohn 1999: 9 n. 10. If HT_f is in $\text{dom } f$, then there is a membership chain from HT_f to itself, contradicting Zermelo-Fraenkel set theory (ZF). Indeed: HT_f is then a member of the ordered pair $\langle HT_f, f(HT_f) \rangle$, which is a member of the function f , which is a member of the proposition HT_f —assuming that HT_f is (inter alia) about f and that a proposition is a set having among its members the objects the proposition is about. Proponents of NP might respond by rejecting the latter assumption or the foundation axiom of ZF (cf. Aczel 1988; Barwise & Etchemendy 1987; Enderton 1977). (Thanks to Stephen Yablo here.)

4. A partial vindication of the Principal Principle.

In §2 I argued that GP entails both PP and CP. In §3 I argued that CP entails NP. Taken together, these conclusions entail my second main thesis: GP entails both PP and NP. It seems then that PP and NP are compatible rather than competing (assuming that GP is consistent); in particular, proponents of NP need not reject PP. Why, then, does Lewis say that PP never applies? To see why, compare OP with NP. Supposing that HT_f is admissible with respect to $\langle Ch(A)=f(A) \rangle$, should $Cr(A|HT_f)$ be $f(A)$ or $f(A|HT_f)$? A conflict looms if $f(A) \neq f(A|HT_f)$. Indeed, Lewis argues that, for some propositions F which he calls “undermining futures”, $f(F) \neq f(F|HT_f)$: $f(F)$ is positive but $f(F|HT_f)$ is zero. The details of Lewis’s argument don’t matter here. Moreover, it turns out that his argument is invalid (Vranas 2002). Nevertheless, for the sake of argument let me grant that $f(F) \neq f(F|HT_f)$. Then OP and NP conflict *provided* that HT_f is admissible with respect to $\langle Ch(F)=f(F) \rangle$. Lewis rejects this admissibility assumption, for a reason which I will examine in a moment. But if HT_f is never admissible, then OP is vacuous. This is roughly why Lewis says that PP never applies.

There are two gaps in the above reasoning, however. First: even if OP doesn’t apply to *cases of undermining*, why doesn’t OP (or PP) apply to more mundane cases? Second: given that OP is just a special case of PP, how does the claim that OP is vacuous entail that so is PP? Lewis, as we will see, does have replies to these objections; but I will argue that his replies fail. My third main thesis in this paper is that, even if OP should be discarded as vacuous, PP need not be; even if in the exotic cases of undermining PP doesn’t apply, in ordinary cases it applies exactly.

Let me deal with Lewis’s reply to the second objection first; the objection that, since OP is just a special case of PP, the vacuousness of OP does not entail that of PP. Lewis’s reply is simple: he claims that, in addition to being entailed by PP, OP *entails* PP (1994: 487) and is thus *equivalent* to PP (cf. 1980: 100). My response is also simple: Lewis’s derivation of PP from OP *relies on the assumption that HT_f is admissible*. This should be intuitively clear even without examining the details of the derivation (see footnote 21 and corresponding text): to derive PP from OP one must *apply* OP, so one must assume that HT_f is admissible. So if, as Lewis claims, HT_f is never admissible, then his derivation of PP from OP fails.¹⁴ It follows that PP need not be vacuous even if OP is.

I turn now to Lewis’s reply to the other objection; the objection that, even if OP doesn’t apply to cases of undermining, OP and PP may still apply to more mundane cases. Let’s see first why Lewis thinks that OP doesn’t apply to cases of undermining. Lewis believes that the chances at a possible world are determined by the total (i.e., past, present, and future) history of the world. If so, then the conjunction of a proposition H that completely specifies the past and present with a proposition F that completely specifies the future entails a proposition T_f that completely specifies the present chances.¹⁵ F is (by definition) an undermining future with respect to HT_f only if T_f is incompatible with T_f : if F were to come about, then (given H) it would complete a total history entailing T_f , and would thus “undermine” T_f . But then FHT_f is impossible, so HT_f reveals information about the future: HT_f entails that F will not come about. This is why Lewis thinks that HT_f is not admissible with respect to $\langle Ch(F)=f(F) \rangle$. Lewis argues that the above reasoning is not limited to cases of undermining:

¹⁴ Lewis might be read as calling ‘OP’ just the *consequent* of what I call ‘OP’, and thus as deriving PP from his OP without assuming that HT_f is admissible. This reading, however, does not square with Lewis’s claim that his OP is a “special case” of PP (1994: 487). (Similarly for Hall 1994: 507-8.)

¹⁵ Strictly speaking, the entailment is restricted to what Lewis calls “worlds like ours”; this is basically why Lewis’s argument is invalid (Vranas 2002). I ignore this complication in the text. F , like H , specifies only the arrangement of “local qualities” (footnote 7).

if chancemaking patterns extend into the future, then *any* use of the Principal Principle is fallacious. For *any* proposition that bears information about present chances thereby bears information about future history. (1994: 485)

In response I ask: why is information about the future never admissible? Lewis gives a couple of examples. Let A be the proposition that a given coin will come up heads at the next toss. Given that $Ch(A)$ is 50% and that the coin *will* come up heads, your conditional credence in A should be 100%, not 50%. More interestingly: given that $Ch(A)$ is 50% and that the coin will come up heads in 99 of the next 100 tosses, your conditional credence in A should be 99%, not 50%. So far, so good (let me grant); but why is *no* information about the future ever admissible? (Cf. Thau 1994: 500; Strevens 1995: 551.) Given that $Ch(A)$ is 50% and that the coin will come up heads in 50 of the next 100 tosses, your conditional credence in A should be 50%. Given that $Ch(A)$ is 50% and that the coin will come up heads at a toss *after* the next, your conditional credence in A should still be 50% (assuming the tosses are independent). Lewis needs, but does not give, an argument to the effect that the *specific* information about the future which is provided by information about present chances is never admissible.

Such an argument is proposed by Strevens: “on a frequency account of chance, *all* information about probabilities is information, in part, about future frequencies. So information about probabilities is itself inadmissible” (1995: 554). I reply first that this argument fails even if frequentism is true. Even if the past and present are known, knowledge of present chances need not determine future frequencies: if you just know that in past tosses (of which there have been exactly five) the proportion of heads was 40% and that in all tosses the proportion (i.e., the chance) is 50%, you still don’t know the proportion in future tosses because you don’t know the *number* of future tosses.¹⁶ Admittedly you know that the proportion in future tosses will be *at least* 50%, but this information about future frequencies is not inadmissible.¹⁷ What *would* be inadmissible is the information that the proportion of heads in future tosses will *exceed* 50%, but *this* information you don’t have: the proportion (or rather the limit) can be exactly 50% if the number of future tosses is infinite.¹⁸ In any case, my second and more important reply to Strevens’s argument is that, as Lewis himself explains (1994: 477), frequentism is false. Lewis does acknowledge that in the simplest case his own account of chance reduces to frequentism, but he also claims that in more complicated cases his account leaves frequentism behind altogether (1994: 481).

Can one avoid assuming frequentism and still construct an argument to the effect that the information about the future which is provided by information about present chances is never admissible? Let’s see first what is the specific information about the future which (on Lewis’s view) is provided by the information that the present chance of A is 50%. It’s this: no future F will occur which (in conjunction with the past and present H) would determine a present chance of A other than 50%. One might think then that the information that $Ch(A)$ is 50% excludes certain future frequencies, namely the frequencies specified by those futures F which it excludes. Not so, however: the claim that a

¹⁶ What if you also know that the coin will be tossed just once? (Cf. the numerical example in Lewis 1994: 488.) Even if the *conjunction* of this information about the future with the information in the text is inadmissible, it doesn’t follow that so is the information in the text alone.

¹⁷ One might grant that this information is not inadmissible but claim that it is not admissible either (see end of footnote 5): given that the proportion of heads in future tosses will be at least 50%, your conditional credence in A (heads at the next toss) *may* but *need not* be 50% if frequentism is true. Although this conclusion is not Strevens’s (he doesn’t make my inadmissible/not admissible distinction), it might suffice to establish that PP does not apply; this is why I go on to say in the text that my second reply to Strevens’s argument is more important than the first.

¹⁸ But wouldn’t a coin disintegrate after sufficiently many tosses? Probably, but arguably the relevant future tosses are not just those of a single coin; they are those of all qualitatively identical coins. Lewis, for example, tries to relate the decay chance of a tritium atom to the decay times of all tritium atoms that ever exist (1994: 477, 482).

future F in which the frequency of heads is (e.g.) 70% will *not* occur is compatible with the claim that a different future F' in which the frequency of heads is also 70% will occur and thus does not exclude a 70% future frequency of heads. So why would the information that F will not occur fail to be admissible? As another example, if F specifies that all future tosses of your coin will come up heads, then given that F will not occur and that $Ch(A)$ is 50% (and given also H) your conditional credence in A (heads at the next toss) should still be 50%: what *else* could it rationally be? One might argue that it should be *less* than 50% because *fewer* futures in which A is true are compatible with H in the presence than in the absence of the information that F will not occur. This argument fails if in both cases *infinitely* many futures are compatible with H . Moreover, why should your conditional credence track the number of compatible futures? It won't do to claim that you should treat all these futures as equally likely: such "principles of indifference" are notoriously problematic (van Fraassen 1989: chap. 12). So again I don't see why the information that F will not occur would fail to be admissible. In sum, even if PP does not apply to cases of undermining, we have been given no good reason to believe that PP does not apply to ordinary—i.e., non-undermining—cases.¹⁹

Doesn't my view introduce an artificially sharp dichotomy between cases to which PP does not apply and cases to which it does? If PP does not apply to (a case involving an) undermining future F , why should PP apply to a proposition A which differs from F only in that it *almost* completely specifies the future? (Cf. Hofer 1997: 331-3.) Because, according to Lewis, "present chances supervene upon the *whole* of history, future as well as present and past" (1994: 482, emphasis added). On such a view HF determines *all* present chances but HA need not determine *any* of them. Isn't then such a view just implausible? Maybe, but if the view is rejected then it turns out that Lewis's derivation of the existence of undermining futures is blocked. So if the view is accepted, then its sharpness accounts for the sharpness of the dichotomy; and if the view is rejected, then Lewis's argument fails to restrict the applicability of PP and the dichotomy disappears.²⁰

Lewis, however, has one more argument up his sleeve. He claims that NP can be used to compute $Cr(A|E<Ch(A)=x>)$ for a generic proposition A (assuming that E is admissible with respect to $<Ch(A)=x>$) and that—contrary to PP—the result of the computation need not be x . Here is why. First, under certain assumptions it can be shown that:

$$(1) Cr(A|E<Ch(A)=x>) \text{ should be } \sum Cr(A|HT_f)Cr(HT_f) / \sum Cr(HT_f).$$

The two sums are over all propositions of the form HT_f which are compatible with E and for which $f(A) = x$.²¹ Now if every HT_f in the numerator sum is admissible with respect to $<Ch(A)=f(A)>$, then

¹⁹ Hofer (1997) and Roberts (2001) propose similar restrictions on the applicability of PP, without however addressing Lewis's arguments to the effect that PP never applies. One might argue that my restriction makes PP useless: we don't know *when* PP applies because we don't know *which* specific propositions are undermining futures. I reply that every ordinary proposition does not completely specify the future and is thus by definition not an undermining future.

²⁰ Lewis's derivation of the existence of undermining futures hinges on excluding the possibility that *every* future, in conjunction with the past and present, determines the *same* present chances (1994: 482). This possibility is not excluded if present chances don't supervene on the whole of history, because then they may supervene on the past and present alone. But the existence of undermining futures follows also from an intermediate view according to which present chances supervene neither on the whole of history nor on the past and present alone but rather on the past and present together with *some* of the future. This intermediate view is blurry, and the corresponding dichotomy between cases to which PP does not apply and cases to which it does is also blurry; but the blurriness of the former would account for the blurriness of the latter.

²¹ Here is a proof of (1), based on Lewis (1980: 99-100; 1994: 489 n. 10) and Hall (1994: 506-8). Suppose that E is admissible with respect to $<Ch(A)=x>$ and that this is so only if E is a disjunction of propositions of the form HT_f (because only propositions about the past and/or present chances are admissible). Given that distinct HT_f are incompatible, $Cr(A|E<Ch(A)=x>) = Cr(AE<Ch(A)=x>)/Cr(E<Ch(A)=x>) = \sum Cr(AHT_f<Ch(A)=x>)/\sum Cr(HT_f<Ch(A)=x>)$ —the two sums being over all HT_f compatible with E . If $f(A) \neq x$, then T_f is incompatible with $<Ch(A)=x>$, so

according to OP $Cr(A|HT_f)$ should be $f(A)$, namely x , so the ratio of the two sums simplifies to x . (This is how OP supposedly entails PP, and why the derivation assumes that HT_f is admissible.) According to NP, however, $Cr(A|HT_f)$ should be $f(A|HT_f)$, and thus apparently need not be x . Lewis comments: “Sad to say, the ... expression derived from (NP) remains unsimplified” (1994: 489 n. 10).

Here is my reply. There is *indirect* reason to believe that the expression derived from NP does simplify to x . A variant of (1) holds in which the left-hand side is just $Cr(A|<Ch(A)=x>)$. This conditional credence should be x according to MP. So if the ratio of the two sums does not simplify to x , then MP contradicts NP. But then CP is inconsistent, since CP entails both MP and NP. So if CP is consistent then the ratio of the two sums simplifies to x .²² This argument is admittedly inconclusive: CP may be inconsistent after all. But given that CP certainly *looks* consistent, the argument shifts the burden of proof to those who claim that the ratio is unsimplifiable. In response they might propose a reductio of my indirect argument: if MP suggests that for any A the ratio of the two sums simplifies to x , then this should be so even when A is replaced with an undermining future—clearly the wrong result. No: this is the *right* result. Even if PP does not apply to cases of undermining, MP does. Even if $Cr(F|H<Ch(F)=x>)$ should be zero, $Cr(F|<Ch(F)=x>)$ should be x : it’s $HF<Ch(F)=x>$, not $F<Ch(F)=x>$, that’s supposed to be impossible when F (given H) undermines $<Ch(F)=x>$ (so that HF entails $<Ch(F)\neq x>$). (So if PP does not apply it’s because H is not admissible with respect to $<Ch(F)=x>$; $<Ch(F)=x>$ itself *is* admissible.) I conclude again that we have been given no good reason to believe that PP does not apply to ordinary cases. This conclusion provides a vindication of PP. This vindication is only partial because for the sake of argument I granted that PP does not apply to cases of undermining. But if Vranas (2002) is right that PP applies even there, then the vindication of PP is considerably strengthened.

5. The indispensability of admissibility.

My judgments of admissibility in the previous sections were based on intuition. The sad fact is that no adequate substantive account of admissibility exists in the literature.²³ So it may be interesting to note that by using CP one can derive a *sufficient* condition for admissibility:

(2) If $Cr(<Ch(A|E)=Ch(A)>)$ should be one, then E is admissible with respect to $<Ch(A)=x>$.

Indeed: if $Cr(<Ch(A|E)=Ch(A)>)$ should be one, then (cf. footnote 6) $Cr(A|E<Ch(A)=x>)$ should be $Cr(A|E<Ch(A|E)=x>)$ and should thus be x by CP. I am not proposing the above condition as *necessary*, however, so I am not offering a *full* substantive account of admissibility. Given the lack of such an account, it would be nice if we could dispense with the notion of admissibility altogether. Hall (1994) thinks that we can, because NP is formulated without reference to admissibility. But this is not so for PP and GP, and these two principles apparently cannot be derived from NP.²⁴ So even if NP dispenses with admissibility, it doesn’t follow that admissibility is ultimately dispensable.

$Cr(AHT_f<Ch(A)=x>)$ and $Cr(HT_f<Ch(A)=x>)$ should be zero. On the other hand, if $f(A) = x$, then T_f entails $<Ch(A)=x>$, so $Cr(AHT_f<Ch(A)=x>)$ should be $Cr(AHT_f)$ —i.e., $Cr(A|HT_f)Cr(HT_f)$ —and $Cr(HT_f<Ch(A)=x>)$ should be $Cr(HT_f)$. (1) quickly follows.

²² I am talking here about the ratio of the two sums in the *variant* of (1) in which the left-hand side is just $Cr(A|<Ch(A)=x>)$. The point is that if *this* ratio simplifies to x then there is reason to believe that the ratio in (1) also does.

²³ For accounts of admissibility see: Lewis 1980: 92-6, 1994: 486-8; Thau 1994: 499-501; Hall 1994: 507-10; Strevens 1995, 1999: 245-6; and some of the references in Kyburg 1981: 773.

²⁴ But didn’t I argue in §4 that that the expression derived from NP does simplify to x ? I did, but this is not enough to derive PP from NP, because the proof of (1) relied on the false assumption that no proposition about the future is ever admissible (footnote 21).

It can be plausibly argued, moreover, that admissibility is indispensable in a deeper sense: *every* chance-credence principle needs an admissibility clause, including those principles (namely MP, CP, and NP) that I formulated without such a clause. Here is why. Assume you know (for sure) that $Ch(A)$ is 50%. Does it follow that your credence in A should be 50%? Given MP as formulated in §2 it does; but in fact it doesn't, because you may *also* know some inadmissible proposition (cf. Strevens 1999: 245). This problem might be fixed in either of two ways. First, by assuming that you have no information inadmissible with respect to $\langle Ch(A)=50\% \rangle$; call this the “no inadmissible information” (NII) clause. Second, by assuming that you have no information relevant to A other than that $Ch(A)$ is 50%; call this the “no other relevant information” (NORI) clause. How to make NII or NORI precise is a vexed question,²⁵ but for present purposes the important point is that every chance-credence principle should be understood as incorporating such a clause. (This is one reason why I said in §2 and §3 that my formulations were very rough.²⁶) To see that (*pace* Hall) this is so even for NP, let K be a proposition that you know for sure. Then $Cr(A|HT_f)$ and $Cr(A|KHT_f)$ should be equal. But according to NP these two conditional credences should be respectively $f(A|HT_f)$ and $f(A|KHT_f)$,²⁷ which need not be equal (your knowledge of K does not entail that $f(K)$ should be one). (Cf. Strevens 1995: 557-9.)

If one uses NORI rather than NII to formulate MP, then it seems that MP (similarly for CP and NP) does dispense with admissibility after all: NORI uses irrelevance rather than admissibility. But if irrelevant information is always admissible, then NORI entails NII. Moreover, since relevant but admissible evidence poses no problem, NORI is unduly strong: when one uses NORI, it's its consequence NII that does the real work. So NORI only nominally “dispenses” with admissibility. The moral is that the dependence of chance-credence principles on admissibility is deeper than meets the eye.

6. Conclusion.

I have proposed a reconciliation between PP and NP. Both principles follow from GP. Proponents of PP need not reject NP: NP is a consequence of the highly intuitive CP. Proponents of NP need not reject PP: even if PP does not apply to cases of undermining, to ordinary cases it applies exactly. What we *should* reject, if my arguments succeed, is the false dichotomy between PP and NP. Both principles can be nonvacuously true; we can have our cake and eat it too.

REFERENCES

- Aczel, Peter. 1988. *Non-well-founded Sets*. Center for the Study of Language and Information, Stanford University: Lecture Notes No. 14.
- Barwise, Jon, & John Etchemendy. 1987. *The Liar: An Essay on Truth and Circularity*. New York: Oxford University Press.

²⁵ Requiring *absence of certainty* about relevant or inadmissible propositions is arguably too weak: what if you are justifiably *almost* certain that such a proposition is true? Requiring *absence of belief* in such propositions (i.e., that they not be in dom Cr) is arguably too strong: what if you know that such a proposition is *false*?

²⁶ Another reason is that provision needs to be made for cases in which a proposition A entails a proposition that specifies the chance of A . If, e.g., the sentence “this sentence expresses a proposition that has a chance 30%” expresses a proposition A , then A entails $\langle Ch(A)=30\% \rangle$, so $Cr(A|\langle Ch(A)=50\% \rangle)$ should be zero—whereas it should be 50% according to MP as formulated in §2. MP does not apply to such cases, and similarly for the other principles (except for NP, which avoids the problem *if* it is stipulated that f assigns value zero to every impossible proposition). (Thanks to Frank Arntzenius here.)

²⁷ *Proof* (cf. Hall 1994: 510). $Cr(A|KHT_f)$ should be $Cr(AK|HT_f)/Cr(K|HT_f)$ and thus by NP should be $f(AK|HT_f)/f(K|HT_f)$ —i.e., $f(A|KHT_f)$.

- Bub, Jeffrey, & Michael Radner. 1968. Miller's paradox of information. *British Journal for the Philosophy of Science* **19**: 63-67.
- Domotor, Zoltan. 1981. Higher order probabilities. *Philosophical Studies* **40**: 31-46.
- Enderton, Herbert. 1977. *Elements of Set Theory*. San Diego: Academic Press.
- Good, Irving. 1970. A suggested resolution of Miller's paradox. *British Journal for the Philosophy of Science* **21**: 288-289.
- Hájek, Alan. 2002. What conditional probability could not be. *Synthese* (forthcoming).
- Hall, Ned. 1994. Correcting the guide to objective chance. *Mind* **103**: 505-517.
- Hofer, Carl. 1997. On Lewis's objective chance: "Humean supervenience debugged". *Mind* **106**: 321-334.
- Howson, Colin, & Graham Oddie. 1979. Miller's so-called paradox of information. *British Journal for the Philosophy of Science* **30**: 253-278.
- Howson, Colin, & Peter Urbach. 1993. *Scientific Reasoning: The Bayesian Approach* (2nd ed.). Chicago: Open Court.
- Ismael, Jenann. 1996. What chances could not be. *British Journal for the Philosophy of Science* **47**: 79-91.
- Jeffrey, Richard. 1970. Review of articles by David Miller *et al.* *Journal of Symbolic Logic* **35**: 124-127.
- Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Kyburg Jr., Henry. 1981. Principle investigation. *The Journal of Philosophy* **78**: 772-778.
- Lewis, David. 1980. A subjectivist's guide to objective chance. Reprinted with postscripts in Lewis 1986.
1986. *Philosophical Papers: Volume II*. New York: Oxford University Press.
1994. Humean Supervenience debugged. *Mind* **103**: 473-490.
- Mackie, John. 1966. Miller's so-called paradox of information. *British Journal for the Philosophy of Science* **17**: 144-147.
- Miller, David. 1966a. A paradox of information. *British Journal for the Philosophy of Science* **17**: 59-61.
1966b. On a so-called so-called paradox: A reply to Professor J. L. Mackie. *British Journal for the Philosophy of Science* **17**: 147-149.
1968. The straight and narrow rule of induction: A reply to Dr Bub and Mr Radner. *British Journal for the Philosophy of Science*. **19**: 145-151.
- Popper, Karl. 1966a. A comment on Miller's new paradox of information. *British Journal for the Philosophy of Science* **17**: 61-69.
1966b. A paradox of zero information. *British Journal for the Philosophy of Science* **17**: 141-143.
- Roberts, John. 2001. Undermining undermined: Why Humean Supervenience never needed to be debugged (even if it's a necessary truth). *Philosophy of Science* **68** (Proceedings): S98-S108.
- Rozeboom, William. 1968. New mysteries for old: The transfiguration of Miller's paradox. *British Journal for the Philosophy of Science* **19**: 345-353.
- Skyrms, Brian. 1988. Conditional chance. In James Fetzer (Ed.), *Probability and Causality: Essays in Honor of Wesley C. Salmon* (pp. 161-178). Dordrecht: Reidel.
- Spohn, Wolfgang. 1999. Lewis' Principal Principle is a special case of van Fraassen's Reflection Principle. Unpublished translation of "Lewis' Principal Principle ist ein Spezialfall von van Fraassens Reflexion Principle". In J. Nida-Rümelin (Ed.), *Rationalität, Realismus, Revision. Proceedings des 3. Kongresses der Gesellschaft für Analytische Philosophie* (pp. 164-173). Berlin: de Gruyter.
- Strevens, Michael. 1995. A closer look at the 'new' principle. *British Journal for the Philosophy of Science* **46**: 545-561.
1999. Objective probability as a guide to the world. *Philosophical Studies* **95**: 243-275.
- Thau, Michael. 1994. Undermining and admissibility. *Mind* **103**: 491-503.
- van Fraassen, Bas. 1980. A temporal framework for conditionals and chance. *The Philosophical Review* **89**: 91-108.
1989. *Laws and Symmetry*. Oxford: Clarendon Press.
- Vranas, Peter. 2002. Who's afraid of undermining? Why the Principal Principle might not contradict Humean Supervenience. *Erkenntnis* (forthcoming).