

Reply to John Hawthorne and Maria Lasonen-Aarnio

(forthcoming in P. Greenough and D. Pritchard, eds., *Williamson on Knowledge*, Oxford University Press)

Timothy Williamson

1. As John Hawthorne and Maria Lasonen-Aarnio appreciate, some of the central issues raised in their 'Knowledge and Objective Chance' arise for all but the most extreme theories of knowledge. In a wide range of cases, according to very plausible everyday judgments, we know something about the future, even though, according to quantum mechanics, our belief has a small but nonzero chance (objective probability) of being untrue. In easily constructed examples, we are in that position simultaneously with respect to many different propositions about the future that are equiprobable and probabilistically independent of each other, at least to a reasonable approximation. Taking the contents of all these pieces of knowledge as premises, we can competently deduce their conjunction, and believe it on that basis. By a very plausible multi-premise closure principle for knowledge, we thereby come to know the conjunction. Since the chance that our belief in the conjunction is true is the product of the chances that the separate conjuncts are true, given independence, it can be made arbitrarily close to zero by choosing an example with enough conjuncts. But this contradicts the very plausible-

sounding principle that a belief constitutes knowledge only if it has a reasonable chance — at any rate over $\frac{1}{2}$ — of being true. Thus if we follow our inclination to accept each of these very plausible claims about knowledge, we are led into inconsistency.

Extreme sceptics happily deny that we know the conjuncts. Eccentrics who regard true belief as sufficient for knowledge happily assert that we know the conjunction. For the rest of us, the problem is more serious.¹ Of course, the argument needs clarification and qualification in various respects, but they can be provided; Hawthorne and Lasonen-Aarnio supply many of the details.

Those who deny the principle of bivalence for future contingents may think that an objective chance is the closest a future contingent can now come to a truth-value. On that view there is no knowledge of chancy future contingents, since knowledge requires truth, and therefore no knowledge of the improbable conjunction. But, by the same reasoning, there is no knowledge of the probable conjuncts, so Hawthorne and Lasonen-Aarnio's challenge is a non-starter. Like them, I happily assume bivalence for future contingents.

The problem is in any case robust. It does not really depend on any contentious metaphysics of the future, for it arises even for knowledge after the critical events have occurred but before the outcomes have been observed or reported. An analogous problem arises in a version of the preface paradox. A meticulous historian writes a long book full of separate factual claims. Given human fallibility, it is almost inevitable that the book will contain errors somewhere or other, for any of which she apologizes in the preface. Nevertheless, she competently deduces the conjunction of all the separate claims in the book (excluding the preface) from its conjuncts and believes it on that basis. As it

happens, she does in fact know each conjunct. Therefore, by closure, she knows the conjunction. But how can she know it when it is almost certain to be false? These analogues may not involve chance in the strict sense, since they concern knowledge of the past, not of the future. Despite the consequent loss of drama, they provoke a very similar epistemological unease: how can the closure principle for knowledge be reconciled with the combination of small probabilities of error for each premise into a large probability of error for the conclusion?

I have just posed the problem without mentioning the epistemological framework distinctive of *Knowledge and its Limits*. Thus the question naturally arises, how far Hawthorne and Lasonen-Aarnio's discussion really concerns a special problem for the theory in *Knowledge and its Limits*, as opposed to the special form that a problem for almost everyone takes when translated into the language of the book. Section 2 explores the general problem of closure in terms of evidential probability, as characterized in the book. Section 3 discusses the divergence between evidential and objective probability. Section 4 considers whether the upshot undermines what the book says about safety and danger.

2. Let $\mathbf{c}_1 \& \dots \& \mathbf{c}_n$ be the conjunction of n equiprobable, mutually probabilistically independent conjuncts $\mathbf{c}_1, \dots, \mathbf{c}_n$. Suppose that, for each i , I know \mathbf{c}_i without knowing that I know \mathbf{c}_i (this sort of possibility is defended at 114-23). More specifically, suppose that, for each i , although I know \mathbf{c}_i , the probability on my evidence that I know \mathbf{c}_i is high but less than 1. On the theory of evidential probability the book defends, with the equation $E = K$ of one's total evidence with one's total knowledge, that I know \mathbf{c}_i entails that the

probability of c_i on my evidence is 1, for c_i is then part of my evidence, and so part of what my evidential probabilities are conditionalized on. That the probability on my evidence that I know c_i is less than 1 entails that I do not know that I know c_i , for otherwise the proposition that I know c_i would be part of my evidence, and so would have probability 1 on my evidence. Conversely, given the regularity assumption that all possibilities consistent with my evidence have nonzero probability on my evidence (225), that it is consistent with what I know that I do not know c_i entails that the probability on my evidence that I know c_i is less than 1. Suppose, furthermore, that the propositions that I know c_i ($1 \leq i \leq n$) are also equiprobable and mutually probabilistically independent on my evidence. That evidence does not favour some of them over others, and does not treat patterns of knowledge and ignorance in some cases as positively or negatively correlated with patterns of knowledge and ignorance in others, but simply as dependent on independent contingencies of the subject matter (this is a simplifying idealization; a version of the argument below still holds on the more realistic assumption that, conditional on knowledge of some c_i , knowledge of others is slightly more probable). Then, although the probability on my evidence of $c_1 \&\dots\& c_n$ is 1, the probability on my evidence that I know every c_i separately is the probability that I know a given c_i raised to the power of n , and so becomes arbitrarily small as n becomes arbitrarily large. Suppose that it is quite clear on my evidence that I believe $c_1 \&\dots\& c_n$ by competent deduction from its conjuncts, and that I do not know $c_1 \&\dots\& c_n$ in any other way, so that I know $c_1 \&\dots\& c_n$ only if I know every c_i separately. Then the probability on my evidence that I know $c_1 \&\dots\& c_n$ is no greater than the probability on my evidence that I know every c_i separately. Thus the probability on my evidence that I know $c_1 \&\dots\& c_n$ becomes

arbitrarily small as n becomes arbitrarily large. Nevertheless, I do in fact know $c_1 \&\dots\&c_n$, by the closure principle, because I do in fact know each of its conjuncts and believe the conjunction by competent deduction from its conjuncts. Hence, even though closure holds, if my judgments of whether I know go with whether it is probable or improbable on my evidence that I know, then I shall judge truly of each conjunct that I know it while also judging falsely that I do not know the conjunction. This suggests that we may be able to explain why such cases appear to be counterexamples to closure, even though really they are not.

We can construct a toy model of epistemic logic to check the coherence of the foregoing account. For worlds we use n -tuples of numbers drawn from the set $\{0, 1, \dots, 2k\}$, where k is a large natural number. Thus there are $(2k+1)^n$ worlds in the model. Think of the n components of a world as its positions on n independent dimensions of a state space; the i th dimension is the one relevant to c_i . Some notation will be convenient: the i th component of the n -tuple w is w_i ; the world just like w except that its i th component is m is $w[i|m]$, so $w[i|m]_i = m$ and $w[i|m]_j = w_j$ if $i \neq j$. Let c_i be true at w if and only if $w_i > 0$. To evaluate knowledge ascriptions in the model, we need an accessibility relation between worlds: as usual, \mathbf{Kp} ('one knows \mathbf{p} ') is true at a world w if and only if \mathbf{p} is true at every world accessible from w . This semantic clause validates the strongest form of closure, on which one automatically knows every conclusion that follows from premises that one knows, and *a fortiori* validates more realistic closure principles; for present purposes, logical omniscience is a harmless idealization. Let x be accessible from w (wRx) if and only if for all i , $|w_i - x_i| \leq k$, that is, w and x do not differ by 'too much' in any of their respective components. In effect, a 'safety' condition is applied to each of the n

dimensions separately. This accessibility relation is obviously reflexive and symmetric. We can easily check that for any world w , \mathbf{c}_i is known (\mathbf{Kc}_i is true) at w if and only if $w_i > k$. For if $w_i > k$ and wRx then $|w_i - x_i| \leq k$, so $x_i > 0$, so \mathbf{c}_i is true at x ; thus \mathbf{Kc}_i is true at w . Conversely, if $w_i \leq k$ then $wRw[i|0]$, because $|w_i - w[i|0]_i| = |w_i - 0| = w_i \leq k$ and if $i \neq j$ then $|w_j - w[i|0]_j| = 0$; but \mathbf{c}_i is false at $w[i|0]$ because $w[i|0]_i = 0$, so \mathbf{Kc}_i is false at w . By a similar argument, for any world w , \mathbf{c}_i is known to be known (\mathbf{KKc}_i is true) at w if and only if $w_i > 2k$; in other words, \mathbf{c}_i is not known to be known (\mathbf{KKc}_i is not true) at any world in this model. In particular, at the world $\langle 2k, \dots, 2k \rangle$, each \mathbf{c}_i is known and none is known to be known.

When evaluating probabilities over the model, it is convenient to assign them to propositions regarded as sets of worlds. In accordance with the approach of *Knowledge and its Limits*, we start with a prior probability distribution Pr . We treat all worlds as initially equiprobable; thus $\text{Pr}(\{w\}) = 1/(2k+1)^n$. The evidence at w is equated with what is known at w , which consists of exactly the set of accessible worlds, $\{x: wRx\}$, since that proposition and nothing stronger is known at w . The probability at w of a proposition \mathbf{p} on the evidence is $\text{Pr}_w(\mathbf{p})$. It results from conditionalizing the prior probability on the evidence at w :

$$\text{Pr}_w(\mathbf{p}) = \text{Pr}(\mathbf{p} \mid \{x: wRx\}) = \text{Pr}(\mathbf{p} \wedge \{x: wRx\}) / \text{Pr}(\{x: wRx\})$$

The ratio is well-defined because R is reflexive, so $\{x: wRx\}$ is nonempty, so $\text{Pr}(\{x: wRx\}) > 0$.

We must check that the model verifies the required probabilistic independence of the n dimensions. To be more precise, for any given i , a proposition \mathbf{p} is *i-based* if and only if for all worlds x and y , if $x_i = y_i$ then \mathbf{p} is true at x if and only if \mathbf{p} is true at y ($1 \leq i$

$\leq n$). That is, whether an i -based proposition is true at a world depends only on the i th component of that world. In particular, \mathbf{c}_i is an i -based proposition. Obviously, the negation of any i -based proposition is also i -based, as is any conjunction of i -based propositions. We can also prove that whenever \mathbf{p} is an i -based proposition, so is \mathbf{Kp} .² Thus \mathbf{Kc}_i and \mathbf{KKc}_i are also i -based propositions. Then we can prove that whenever for each i \mathbf{p}_i is an i -based proposition, $\mathbf{p}_1, \dots, \mathbf{p}_n$ are mutually probabilistically independent on the evidence in any world, in the usual sense that the probability (on the evidence at that world) of their conjunction is the product of the probabilities (on the evidence at that world) of the conjuncts.³ Although a model could have been constructed in which the evidence at some worlds establishes epistemic interdependences between the different dimensions, for present purposes we can do without such complications. In particular, $\mathbf{c}_1, \dots, \mathbf{c}_n$ are mutually probabilistically independent on the evidence in any world. Thus the epistemic propositions $\mathbf{Kc}_1, \dots, \mathbf{Kc}_n$ are also mutually probabilistically independent on the evidence in any world. But, on the evidence in the world $\langle 2k, \dots, 2k \rangle$, for any given i , the probability that \mathbf{c}_i is known is $k/(k+1)$.⁴ By probabilistic independence, the probability of the conjunction $\mathbf{Kc}_1 \ \&\dots\ \& \ \mathbf{Kc}_n$ is $(k/(k+1))^n$. That is the probability that each conjunct is known. But, by the closure principle built into the model, knowing a conjunction ($\mathbf{K}(\mathbf{c}_1 \ \&\dots\ \& \ \mathbf{c}_n)$) is equivalent to knowing all the conjuncts ($\mathbf{Kc}_1 \ \&\dots\ \& \ \mathbf{Kc}_n$). Thus the probability on the evidence in $\langle 2k, \dots, 2k \rangle$ that the conjunction $\mathbf{c}_1 \ \&\dots\ \& \ \mathbf{c}_n$ is known is also $(k/(k+1))^n$. For fixed k , this probability becomes arbitrarily close to 0 as n becomes arbitrarily large. Thus, for suitable k and n , the world $\langle 2k, \dots, 2k \rangle$ exemplifies just the situation informally sketched: for each conjunct one knows it without knowing that one knows it, and it is almost but not quite certain on

one's evidence that one knows the conjunct; one also knows the conjunction without knowing that one knows it, and it is almost but not quite certain on one's evidence that one does *not* know the conjunction.

In some examples, one's epistemic position with respect to each conjunct is better: one not only knows it but knows that one knows it. If one also knows the relevant closure principle, and knows that one satisfies the conditions for its application, one may even know that one knows the conjunction. Consequently, the probability on one's evidence that one knows the conjunction is 1. However, the previous pattern may still be repeated at a higher level of iterations of knowledge. For example, for each conjunct one knows that one knows it without knowing that one knows that one knows it, and it is almost but not quite certain on one's evidence that one knows that one knows the conjunct; one also knows that one knows the conjunction without knowing that one knows that one knows it, and it is almost but not quite certain on one's evidence that one does *not* know that one knows the conjunction. To adapt the previous model to this case, we can simply expand the set of worlds by using n -tuples of numbers from the set $\{0, 1, \dots, 3k\}$ rather than $\{0, 1, \dots, 2k\}$, leaving the definitions of accessibility and the truth-conditions of the c_i unchanged (so c_i is true at w if and only if $w_i > 0$); then $\langle 3k, \dots, 3k \rangle$ is a world of the required type. More generally, if one uses as worlds n -tuples of numbers from the set $\{0, 1, \dots, hk\}$, leaving the other features of the model unchanged, then $\langle hk, \dots, hk \rangle$ will be a world at which one has $h-1$ but not h iterations of knowledge of each conjunct, and it is almost but not quite certain on one's evidence that one has $h-1$ iterations of knowledge of the conjunct; one also has $h-1$ but not h iterations of

knowledge of the conjunction, and it is almost but not quite certain on one's evidence that one does *not* have $h-1$ iterations of knowledge of the conjunction.

Many other variations can be played on the same theme. The general idea is this. One attains a given epistemic status E with respect to each conjunct, without knowing that one does (this is possible by the anti-luminosity argument). By a principle of multi-premise closure for E , one also attains status E with respect to the conjunction (supposing E to be an epistemic status of a type to which multi-premise closure considerations apply), without knowing that one does. Then for each conjunct it may be almost certain on one's evidence that one attains E with respect to it, even though it is almost certain on one's evidence that one does *not* attain E with respect to the conjunction. Hence multi-premise closure may appear to fail for E even though it really holds, for if one's judgments of whether one attains E with respect to a given proposition go with whether it is probable or improbable on one's evidence that one attains E with respect to that proposition, then one will judge that one attains E with respect to each conjunct but not with respect to the conjunction.⁵

General principles similar to closure are often more secure than apparent counterexamples to them. Consider a loose analogy. Suppose that my way of judging tallness perceptually has the effect that I am more likely to judge a thin person tall than a fat person of the same or slightly greater height. Fat y is slightly taller than thin x . In a given context in which I am looking at neither x nor y , I may simultaneously have both these dispositions:

- (a) on looking at x alone, to judge that he is tall;
- (b) on looking at y alone, to judge that he is not tall;

In the same context I may well also have the disposition on looking at x and y together, to judge that y is taller than x . This would not be plausible as a counterexample to the general monotonicity principle that if x is tall and y is at least as tall as x then y is tall. Rather, we should hold on to the general principle and conclude that my dispositions to judge in particular cases whether people are tall are not wholly accurate. Similarly, concerning knowledge, we should hold on to the general principle of closure and conclude that our dispositions to judge in preface-like cases whether people know are not wholly accurate.

We cannot plausibly resolve these problem cases by postulating contextual variation in the reference of ‘know’ or ‘tall’. It does not help to say that the reference of ‘tall’ varies between the context in which I am looking at x alone and the context in which I am looking at y alone (with ‘tall’ as used in the former context applying to both x and y and as used in the latter applying to neither), for the problem concerns the extension of ‘tall’ as used in the single *original* context in which I have both (a) and (b) as unmanifested dispositions. If that extension is closed under the monotonicity principle, it does not perfectly match those dispositions. Similarly, concerning knowledge, the problem concerns the extension of ‘know’ as used in a single everyday context in which we have both (a*) and (b*) as unmanifested dispositions, with respect to a preface-like case in which the subject clearly satisfies the conditions for a suitable version of the principle that knowledge is closed under competent deduction to apply:

(a*) on considering any conjunct, to judge that the subject knows it;

(b*) on considering the conjunction, to judge that the subject does not know it.

If that extension is closed under the deductive closure principle, it does not perfectly match those dispositions. We have seen in this section how such mismatches can arise, without denying closure or falling into scepticism.

3. Hawthorne and Lasonen-Aarnio formulate their challenge in terms of objective chance, not evidential probability. Much of its force will be felt by any view that endorses a plausible closure principle for knowledge and is robustly anti-sceptical about knowledge of the future. For the view will imply that one can know a long conjunction about the future, even though there is a high objective chance that one's belief is false. Even when it is luckily true, won't that be a Gettier case rather than a case of knowledge? Similar problems arise for knowledge not of future contingents, as in the preface paradox. Although Hawthorne and Lasonen-Aarnio show that this challenge requires various refinements, as in their Low Chance principle, the underlying problem remains.

On the account defended here and in the book, knowledge corresponds to the highest evidential probability. Before considering the relation between knowledge and chance, it is therefore worth asking a more general question: what is the relation between evidential probability and chance? The answer is: very little. In particular, zero chance is compatible with any level of probability on one's evidence. For example, suppose that you know that a given fair coin was tossed n times in the past, and that the tosses were independent, but have no further information about the outcomes. On any reasonable view, the probability on your evidence that not every toss came up heads is $(2^n - 1)/2^n$, and so becomes arbitrarily close to 1 as n becomes arbitrarily large. Indeed, if n is countably infinite, the probability is 1 (by the argument of Williamson (2007c), with the order of

the tosses reversed). But if by chance every toss did come up heads, then the chance that not every toss came up heads is now 0. Thus not even chance 0 for a proposition puts a non-trivial upper bound on its evidential probability. The same point applies to the future. Suppose that we know only that a coin has already been selected for a future toss, and that with probability x on our evidence a two-headed coin was selected, otherwise a two-tailed coin was selected. Then the probability on our evidence that the selected coin, whichever it is, will come up heads is x . But if in fact the two-tailed coin was selected, the chance that the selected coin will come up heads is 0. Thus if low chance puts a non-trivial bound on knowledge, that is a very specific feature of knowledge; it does not reflect a more general correlation between chance and evidential probability.

Of course, if one knows something about chances, that knowledge will contribute to one's evidence, and thereby to probabilities on one's evidence. In examples of Hawthorne and Lasonen-Aarnio's kind, such knowledge is often available. One can know that the long conjunction is objectively improbable, and that each of its conjuncts is objectively probable. Indeed, for simplicity, we may even pretend that the exact chances of the long conjunction and of each of its conjuncts are known.

A salient proposal about the impact of evidence about chances comes from David Lewis's Principal Principle:

Let C be any reasonable initial credence function. Let t be any time. Let x be any real number in the unit interval. Let \mathbf{ch} be the proposition that the chance, at time t , of \mathbf{p} 's holding equals x . Let \mathbf{e} be any proposition compatible with \mathbf{ch} that is admissible at t . Then $C(\mathbf{p} \mid \mathbf{ch} \ \& \ \mathbf{e}) = x$.⁶

For suppose that one's evidential probabilities at t result from conditionalizing some reasonable initial credence function on one's knowledge of the chances at t and other evidence admissible at t . Then by the Principal Principle one's evidential probabilities at t for the long conjunction and for each of its conjuncts will have the same value as the respective chances.

As Hawthorne and Lasonen-Aarnio point out, examples of contingent *a priori* knowledge will force some revision of the Principal Principle. However, since those examples turn on the use of special rigidifying devices such as an 'actually' operator that are not directly germane here, it is not obvious that the required qualifications will make any difference for present purposes.

Of more immediate concern is that the Principal Principle itself poses a challenge to knowledge of the future, given the closure of knowledge under competent deduction. For suppose that I know each conjunct about the future and believe their conjunction on the basis of competent deduction; by closure, I know the conjunction. However, I also know their respective chances. Suppose that my credences at t result from conditionalizing some reasonable initial credence function on the known chances at t and other evidence admissible at t . Then by the Principal Principle my credences at t for the conjunction and for each of its conjuncts have the same value as the respective chances. Thus my credence in each conjunct is very high, while my credence in the conjunction is very low. But this seems to violate the principle that knowledge entails belief. How can I know the conjunction if my credence in it is very low? Indeed, given the Principal

Principle, any adjustment of my credences to assign high credence to the conjunction looks irrational.

Lewis himself denies that knowledge entails belief. Following Woozley and Radford, he allows knowledge without belief 'in the case of the timid student who knows the answer but has no confidence that he has it right, and so does not believe what he knows' (1996, 1999: 429). But merely postulating knowledge without belief does not solve the problem. For the problematic cases are not simply ones in which I do not in fact believe the chancy conjunction; they are cases in which I *ought not* to believe the conjunction. How can one know **p** if one is right not to believe **p**?

For Lewis, the solution is not to reject closure. On his account, one knows **p** if and only if **p** holds in every [contextually relevant] possibility left uneliminated by one's evidence, where evidence consists of perceptual experiences and memories and eliminates just those possibilities in which one's perceptual experience and memories have a different content (Lewis 1996, 1999: 422-5). As Lewis realized, this implies a form of logical omniscience, with respect to any fixed context: one's knowledge is closed under necessary consequence, whether or not one carries out the appropriate deductions, for if **q** holds in every possibility in which **p**₁, ..., **p**_{*n*} all hold, and each of **p**₁, ..., **p**_{*n*} holds in every possibility relevant in a given context, then **q** holds in every possibility relevant in that context (1996, 1999: 440-2).⁷ Nor is it consonant with Lewis's methodology to become a sceptic about the future. A relaxed conversational context can make a bizarre future possibility irrelevant in his sense, so that one knows a chancy future contingent. Given a context lax enough to make most such possibilities irrelevant, one may know a future contingent even though it has a very low chance of holding, and one's credence in

it — rationally obtained by updating a reasonable initial credence function on known chances and other admissible evidence — is equally low. This uneasy situation arises because possibilities ignored in determining knowledge are not ignored in determining chance and therefore, by the Principal Principle, in determining rational credence.

Lewis has a Rule of Belief, according to which ‘A possibility that the subject believes to obtain is not properly ignored, whether or not he is right to so believe’. A possibility that is not properly ignored is contextually relevant for the purposes of Lewis’s account of ‘know’. He amplifies the rule to take account of degrees of belief and degrees of specificity in possibilities (1996, 1999: 428):

A possibility may not be properly ignored if the subject gives it, or ought to give it, a degree of belief that is sufficiently high, and high not just because the possibility in question is unspecific.

Thus if the subject, in obedience to the Principal Principle, gives the possibility that the chancy conjunction is false a very high degree of belief, then the Rule of Belief seems to imply that the possibility that the conjunction is false is not properly ignored after all. Of course, in a possibility in which the conjunction is false, one of the conjuncts is false too — say, c_i . But Lewis also has a Rule of Resemblance, according to which if one possibility saliently resembles another, and the former possibility may not properly be ignored (in virtue of rules other than this one), then the latter possibility also may not properly be ignored (1996, 1999: 429). Given the nature of Hawthorne and Lasonen-Aarnio’s examples, for each j the possibility in which c_i is false saliently resembles a

possibility in which c_j is false ($1 \leq j \leq n$). Since the possibility in which c_i is false may not properly be ignored (in virtue of the Rule of Belief), the possibility in which c_j is false may not properly be ignored (in virtue of the Rule of Resemblance). Moreover, all the possibilities in question are uneliminated by the subject's evidence, as Lewis understands it. Therefore, on Lewis's account, none of the conjuncts is known after all: a highly sceptical result. Invoking the Rule of Belief in an attempt to reconcile Lewis's contextualism about knowledge with the Principal Principle merely undoes the anti-sceptical work that the contextualism was designed to do.

An alternative is to permit updating on evidence that is inadmissible in the sense of the Principal Principle. For example, I may update by conditionalizing on some contingent truth that I know about the future. Then I may have credence 1 in that truth, even though I know that its chance is less than 1. That is consistent with the Principal Principle, which is logically neutral as to the results of conditionalizing on inadmissible evidence, despite the forbidding connotations of the word 'inadmissible'. For Lewis, however, future contingents do not constitute evidence, in the sense in which he holds that one knows something if and only if it holds in every relevant possibility left uneliminated by one's evidence. Rather, as already noted, he equates one's evidence with the fact about the present that 'one's entire perceptual experience and memory are just as they actually are' (Lewis 1996, 1999: 424), which is admissible. In discussing the Principal Principle, Lewis works with a more liberal notion of evidence, envisaging any proposition strictly about the past as admissible evidence; but propositions strictly about the future remain the paradigms of inadmissibility.⁸ In contrast, by equating one's total evidence with one's total knowledge, including one's knowledge of the future, my

approach permits one to update by conditionalizing on inadmissible evidence in Lewis's sense. In particular, if I know each of the conjuncts, then their conjunction automatically receives probability 1 on my evidence, because each of its conjuncts does, whether or not I carry out the deduction. This is not to reject the Principal Principle but to move outside its conditions of application. The equation $E = K$ enables one to avoid the combination of knowledge with very low rational credence that threatens to arise on Lewis's view.

Once we have granted that there is some knowledge of future outcomes in addition to knowledge of their present chances, we are not entitled to assume that the latter always screens out the former evidentially. Strange though it may sound, we cannot take for granted that there is no knowledge of future outcomes whose chances are known to be low. *A fortiori*, we cannot take for granted that there is no knowledge of future outcomes whose chances *are* low. Indeed, as Hawthorne and Lasonen-Aarnio show, there can be such knowledge, in cases of the contingent *a priori*. If there is a defensible principle in the vicinity, it is something like their *Low Chance*:

For all worlds w , times t , subjects s , belief-episodes B , and propositions \mathbf{p} , if at t in w s 's belief-episode B expresses proposition \mathbf{p} , at t in w the chance that B expresses a true proposition is low, and at t in w s is not inadmissibly connected to the future, then s does not know \mathbf{p} at t in w .⁹

But not even Low Chance is satisfactory as it stands. For example, if s at t in w has two belief-episodes B and B^* , both of which express \mathbf{p} , where at t in w the chance that B expresses a true proposition is low but the chance that B^* expresses a true proposition is

high, then presumably s may still know \mathbf{p} at t in w . To register this point, we may expand the consequent of Low Chance to ‘ s does not know \mathbf{p} at t in w as far as B goes’. Once all such required qualifications have been added to Low Chance, the result hardly looks self-evident.

4. The question remains whether I am committed to something similar enough to Low Chance to make trouble given considerations in the book, such as the conception of knowledge as safely true belief. In effect, Hawthorne and Lasonen-Aarnio suggest such an argument. One premise is their *High Chance – Close Possibility* principle HCCP.

HCCP

For any world w , time t , and proposition \mathbf{p} , if the chance at t in w of \mathbf{p} is high, then there is a close branching possibility at t in w in which \mathbf{p} holds.

One might take the conception of knowledge as safely true belief to be committed to SAFETY*:

SAFETY*

For all worlds w , times t , subjects s , belief-episodes B , and propositions \mathbf{p} , s knows \mathbf{p} at t in w as far as B goes only if B expresses a true proposition in every possibility that is close at t in w .

From HCCP and SAFETY* we can derive a principle similar to Low Chance:

LC*

For all worlds w , times t , subjects s , belief-episodes B , and propositions \mathbf{p} , if the chance at t in w that B expresses a true proposition is low then s does not know \mathbf{p} at t in w as far as B goes.

For suppose that the chance at t in w that B expresses a true proposition is low. Then the chance at t in w that B does not express a true proposition is high (something has a low chance if and only if its contradictory has a high chance). Therefore, by HCCP, at t in w there is a close possibility in which B does not express a true proposition. Therefore, by SAFETY*, s does not know \mathbf{p} at t in w as far as B goes. Thus HCCP and SAFETY* together entail LC*. In the cases at issue, LC* excludes knowledge of the long conjunction of future contingents.¹⁰

Hawthorne and Lasonen-Aarnio suggest that I am *prima facie* committed to HCCP by the discussion of close possibility in terms of ease and difficulty, safety and danger in the book (123-4): couldn't a given high chance event easily occur? There is no straightforward connection. As emphasized in the book, determinism does not trivialize safety; safety and danger are not defined in terms of chances (123). Although the deterministic laws could not easily be broken, a ball precariously balanced on the tip of a cone is not safe from falling, for the initial conditions could easily have been slightly different.¹¹ This point does not refute HCCP, for presumably in a deterministic world the chance at t in w of \mathbf{p} is high only if \mathbf{p} holds in w , which is therefore itself a close branching possibility at t in w in which \mathbf{p} holds. Nevertheless, a gap between closeness

remains to be bridged: ‘branching’ and ‘close’ are not equivalent. Even in a non-deterministic world, not all close possibilities are branching possibilities. But why shouldn’t most or all branching possibilities be close possibilities? In that case, SAFETY* will still have sceptical consequences.

As Hawthorne and Lasonen-Aarnio note, the discussion in the book is conducted in terms of close subject-centred cases rather than close worlds. It also adverts to one’s basis for a belief, conceived not specifically as one’s warrant or evidence for it but more generally as the epistemically relevant features of the belief. Here is a pertinent formulation of safety (compare 102):

SAFETY

If in a case α one knows \mathbf{p} on a basis b , then in any case close to α in which one believes a proposition \mathbf{p}^* close to \mathbf{p} on a basis close to b , \mathbf{p}^* is true.^{12,13}

Hawthorne and Lasonen-Aarnio suggest that multi-premise closure may fail under such a subject-centred conception of safety. Let us first see how SAFETY is compatible with multi-premise closure.

Say that one *safely believes* \mathbf{p} on a basis b in a case α if and only if one believes \mathbf{p} on basis b in α and in any case close to α in which one believes a proposition \mathbf{p}^* close to \mathbf{p} on a basis close to b , \mathbf{p}^* is true. Thus SAFETY says that knowledge entails safe belief. Is safe belief closed under competent deduction? More specifically, if one safely believes some premises and believes a conclusion on the basis of competent deduction from those

premises as believed on the relevant bases, does it follow that one safely believes the conclusion on that basis?

A positive answer requires some link between closeness with respect to the conclusion and the closeness with respect to the premises. This will do:

DEDUCTION

If in case α one believes a conclusion \mathbf{q} on a basis b , which consists of competent deduction from $\mathbf{p}_1, \dots, \mathbf{p}_n$ as believed on bases b_1, \dots, b_n respectively, then in any case close to α in which one believes a conclusion \mathbf{q}^* close to \mathbf{q} on a basis b^* close to b , b^* consists of a truth-preserving deduction from propositions $\mathbf{p}^*_1, \dots, \mathbf{p}^*_n$ close to $\mathbf{p}_1, \dots, \mathbf{p}_n$ respectively as believed on bases b^*_1, \dots, b^*_n close to b_1, \dots, b_n respectively.

On the natural conception of bases underlying DEDUCTION, the basis of an inferential belief incorporates the bases of the beliefs from which it was inferred; the basis is not merely the inference itself. This cumulative conception of bases will be crucial for multi-premise closure. As for ‘truth-preserving deduction’ in DEDUCTION, it replaces ‘competent deduction’ in SAFETY because cases just above the threshold for competence may be close to cases just below the threshold. ‘Truth-preserving’ stands to ‘competent’ for deduction as ‘true’ stands to ‘known’ for propositions; thus DEDUCTION embodies a safety conception of deductive competence. That is legitimate, for our concern is precisely with the implications of safety for multi-premise closure. Although DEDUCTION may require some further fine-tuning, it will do for present purposes.

To see that DEDUCTION implies that safe belief is closed under competent deduction, suppose that in a case α one safely believes premises $\mathbf{p}_1, \dots, \mathbf{p}_n$ on bases b_1, \dots, b_n respectively and believes a conclusion \mathbf{q} on basis b , which consists of a competent deduction from $\mathbf{p}_1, \dots, \mathbf{p}_n$ as believed on bases b_1, \dots, b_n respectively. Consider any case β close to α in which one believes a conclusion \mathbf{q}^* close to \mathbf{q} on a basis b^* close to b . Therefore, by DEDUCTION, in β one believes \mathbf{q}^* on the basis of truth-preserving deduction from propositions $\mathbf{p}^*_1, \dots, \mathbf{p}^*_n$ close to $\mathbf{p}_1, \dots, \mathbf{p}_n$ respectively as believed on bases b^*_1, \dots, b^*_n close to b_1, \dots, b_n respectively. By definition of safety, since β is close to α and in β one believes $\mathbf{p}^*_1, \dots, \mathbf{p}^*_n$ on bases close to b_1, \dots, b_n respectively, $\mathbf{p}^*_1, \dots, \mathbf{p}^*_n$ are true in β . By DEDUCTION the deduction of \mathbf{q}^* from $\mathbf{p}^*_1, \dots, \mathbf{p}^*_n$ is truth-preserving in β , so \mathbf{q}^* is true in β . This is just what was needed to show that in α one safely believes \mathbf{q} . Thus safe belief is closed under competent deduction. SAFETY, formulated in terms of close cases, is quite compatible with multi-premise closure.

The pooling of knowledge from many individuals by testimony can be similarly explained. When the hearer knows something on the basis of testimony from a knowledgeable speaker, the basis of the hearer's belief incorporates the basis of the speaker's belief, even if the hearer is largely ignorant of the latter basis. Bases need not be introspectively available. Again, in memory-based knowledge, the basis of one's present belief incorporates the basis of one's earlier belief, even if one has forgotten the latter basis.

Does chance create counterexamples to SAFETY? Let α be the actual case in which I drop my marble and know that it will land on the floor. Still, there is a small non-zero chance that it will not land on the floor. Let β be a case just like α up to now in

which I drop my marble but, by a quantum-mechanical blip, it does not land on the floor; since the laws of nature are indeterministic, the same ones can hold in β as in α . In β , I still believe that my marble will land on the floor, but my false belief does not constitute knowledge. Why isn't β a case close to α in which I believe on exactly the same basis that my marble will land on the floor? If so, this is a counterexample to SAFETY. But the occurrence of an event in β that bucks a relevant trend in α may be a relevant lack of closeness between α and β , even though the trend falls well short of being a strict law. Thus β is not close to α after all; perhaps the belief's basis in β is also not close to its basis in α . The accumulation of such cases may then yield violations of HCCP, as formulated in terms of cases rather than possibilities.

This does not mean that safety is totally unconstrained by chance. For example, the pattern of chances may determine whether a general type of causal process by which beliefs are acquired counts as a form of perception, and so as a basic source of knowledge of the environment: if the chances of error are too great, it does not count as perception. That is, chance constrains safety globally, not locally case by case.

The structural divergence between knowledge and high chance that Hawthorne and Lasonen-Aarnio exploit in trying to separate knowledge from safety (as ordinarily conceived) is analogous to a structural divergence between safety (as ordinarily conceived) and high chance. Suppose that I am not safe from being shot. On the ordinary conception, it follows that there is someone x such that I am not safe from being shot by x (assume that if I am shot, I am shot by someone). On the high chance conception of safety, that is a *non sequitur*. For each individual x , the chance of my being shot by x may be low enough for me to count as safe from being shot by x , even though the chance of

my being shot by someone or other may be too high for me to count as safe from being shot. Again, on the ordinary conception of danger, if for each individual x I am in no danger of being shot by x then I am in no danger of being shot. On a chancy conception of danger, even if for each individual x I am in no danger of being shot by x , I may still be in danger of being shot. The ordinary conception posits just the sort of closure principle for safety that the high chance conception undermines. If $\mathbf{p}_1, \dots, \mathbf{p}_n$ entail \mathbf{q} then ‘ \mathbf{p}_1 is safe from falsity’, ..., ‘ \mathbf{p}_n is safe from falsity’ entail ‘ \mathbf{q} is safe from falsity’.

The ordinary conception of safety can look like a primitive refusal to acknowledge the potential for many small risks to add up to a big one. But that is already to misconstrue the ordinary conception. It is a ‘no risk’ conception of safety, not a ‘small risk’ conception. This is achieved not by a confused equation of little with nothing but by a tight restriction to close cases, where closeness is determined by considerations of similarity (to the actual case) as much as of chance. To be safe from a danger is to avoid it in *all* close cases.

Unlike the high chance conception of safety, the ordinary conception of safety at least delivers the elementary but crucial consequence that if one is safe from undergoing something then one does not undergo it: high chance events do not always occur, but the actual case is close to itself. If someone was shot, he was not safe from being shot. This factiveness of safety complements its closure under logical consequence. Together they combine to make the logic of safety at least as strong as the modal logic T (= KT), the weakest normal modal logic with the axiom $\Box \mathbf{p} \rightarrow \mathbf{p}$; it serves as an idealized but non-luminous epistemic logic in *Knowledge and its Limits* (305).

The closeness conception of safety might be thought to be of less practical use than the high chance conception. But in many situations it is the opposite. Often we have so little idea of the objective chances that it is infeasible to reason in terms of them. Rather, we think in terms of making ourselves safe from a disjunction of dangers by making ourselves safe from each disjunct separately. That way of thinking assumes a closure principle for safety that the closeness conception can deliver and the high chance conception cannot. The chance of the disjunction is much higher than the chance of any disjunct, but if each disjunct is avoided in all close cases, so is their disjunction. For that to be achievable in practice, many branching possibilities will have to count as not close. Hawthorne and Lasonen-Aarnio show in effect how far the closeness and high chance conceptions of safety can diverge in some cases, but that is no fault in the closeness conception. Similarly, the wide divergence between them in deterministic worlds enables the closeness conception to do much better there than the high chance conception, which disallows unrealized dangers in those cases.

Of course, in many situations the best way of handling risk does involve reasoning about chances, where their numerical values can be reasonably estimated. Even there, a closeness conception may still be needed in the background, since we can never take account of all the bizarre outcomes that have a slight chance of occurring through quantum-mechanical blips, just as known evidence is in the background of reasoning about probabilities on one's evidence. No concept is a panacea. Sometimes thinking with probabilistic concepts is more useful than thinking with the concept of safety. When we are too much in the dark about the probabilities, thinking with the concept of safety is often more feasible than thinking with probabilistic concepts.

Once we understand the distinctive structural virtues of the closeness conception of safety, we can more easily understand the corresponding virtues of the ordinary conception of knowledge. They depend on the same features, such as closure and factiveness. It is not perverse to focus on a property that is cumulative (closure) and success-oriented (factiveness). Since high chance lacks both those features, the failure of closeness to map neatly onto the probabilistic structure of chances becomes an advantage.

What still requires much more detailed investigation is the nature of closeness, and its relation to past and future similarities. We cannot expect it to be a perfectly natural relation; given the anti-reductionism about knowledge for which *Knowledge and its Limits* argues, we cannot expect to identify just what degree and kind of safety is required for knowledge in non-epistemological terms. Still, we do not want closeness to be *too* unnatural.¹⁴ On the ordinary conception, safety is firmly enough rooted in the actual structure of the world, irrespective of its appearance to the agent (Hawthorne and Lasonen-Aarnio call safety ‘objective’). Knowledge also has such roots. Its divergence from high chance does not prevent it from being as natural and objective as one can reasonably demand of any epistemic matter. The divergence is a price well worth paying for the structural virtues of knowledge.¹⁵

Notes

1 In the book I endorse an example of Michael Slote's that requires that I do not know that I shall not be run over by a bus tomorrow (255). This has been interpreted, not unreasonably, as an endorsement of the general view that we cannot know future contingents. I did not intend anything quite so general. It would have been better to have emphasized philosophers' absent-mindedness when crossing roads. In any case, as noted below, the problem that Hawthorne and Lasonen-Aarnio raise can be generalized far beyond future contingents.

2 Proof: Suppose that \mathbf{p} is i -based and $x_i = y_i$. Suppose also that \mathbf{Kp} is false at x . Then for some z , xRz and \mathbf{p} is false at z . But then $yRy[i|z_i]$, for $|y_i - y[i|z_i]_i| = |y_i - z_i| = |x_i - z_i|$ (because $x_i = y_i$) $\leq k$ (because xRz), and if $i \neq j$ then $|y_j - y[i|z_i]_j| = 0$. Moreover, \mathbf{p} is false at $y[i|z_i]$ because \mathbf{p} is false at z and i -based and $y[i|z_i]_i = z_i$. Hence \mathbf{Kp} is false at y . Thus if \mathbf{Kp} is true at y then \mathbf{Kp} is true at x . The converse follows by parity of reasoning.

3 For any world w , proposition \mathbf{q} and $1 \leq i \leq n$, set:

$$\#(i, \mathbf{q}, w) = \{j: 0 \leq j \leq 2k, \mathbf{q} \text{ is true at } w[i|j] \text{ and } |w_i - j| \leq k\}.$$

For each i , let \mathbf{p}_i be i -based. Note that for any worlds w and x , the following conditions are equivalent:

- (i) wRx and $\mathbf{p}_1 \ \&\dots\ \& \ \mathbf{p}_n$ is true at x
- (ii) for all i , $x_i \in \#(i, \mathbf{p}_i, w)$

The proof is trivial: for each i , \mathbf{p}_i is true at x if and only if it is true at $w[i|x_i]$ since \mathbf{p}_i is i -based. Now let $|X|$ be the cardinality of the set X . Since Pr makes all worlds equiprobable, $\text{Pr}_w(\mathbf{p}_1 \&\dots\& \mathbf{p}_n) = |\{x: wRx \text{ and } \mathbf{p}_1 \&\dots\& \mathbf{p}_n \text{ is true at } x\}|/|\{x: wRx\}|$ for a given world w . By the equivalence of (i) and (ii): $|\{x: wRx \text{ and } \mathbf{p}_1 \&\dots\& \mathbf{p}_n \text{ is true at } x\}| = |\{x: \text{for all } i, x_i \in \#(i, \mathbf{p}_i, w)\}| = |\#(1, \mathbf{p}_1, w)| \dots |\#(n, \mathbf{p}_n, w)|$. By the special case of this equation in which each \mathbf{p}_i is replaced by the tautology \mathbf{t} (which is trivially i -based for any i), $|\{x: wRx\}| = |\#(1, \mathbf{t}, w)| \dots |\#(n, \mathbf{t}, w)|$. Consequently:

$$\text{Pr}_w(\mathbf{p}_1 \&\dots\& \mathbf{p}_n) = (|\#(1, \mathbf{p}_1, w)| \dots |\#(n, \mathbf{p}_n, w)|) / (|\#(1, \mathbf{t}, w)| \dots |\#(n, \mathbf{t}, w)|)$$

For any given i , consider another special case in which \mathbf{p}_j is replaced by \mathbf{t} whenever $i \neq j$.

Since $n-1$ of the ratios cancel out, $\text{Pr}_w(\mathbf{p}_i) = |\#(i, \mathbf{p}_i, w)| / |\#(i, \mathbf{t}, w)|$. Therefore

$$\text{Pr}_w(\mathbf{p}_1 \&\dots\& \mathbf{p}_n) = \text{Pr}_w(\mathbf{p}_1) \dots \text{Pr}_w(\mathbf{p}_n), \text{ as required.}$$

4 Proof: We have already established that \mathbf{Kc}_i is true at a world x if and only if $x_i > k$. Thus, in the notation of fn. 3, $\#(i, \mathbf{Kc}_i, \langle 2k, \dots, 2k \rangle) = \{j: k < j \leq 2k\}$, so

$$|\#(i, \mathbf{Kc}_i, \langle 2k, \dots, 2k \rangle)| = k, \text{ while } |\#(i, \mathbf{t}, \langle 2k, \dots, 2k \rangle)| = \{j: k \leq j \leq 2k\}, \text{ so}$$

$$|\#(i, \mathbf{t}, \langle 2k, \dots, 2k \rangle)| = k+1. \text{ By the formula for } \text{Pr}_w(\mathbf{p}_i) \text{ in fn. 3, } \text{Pr}_{\langle 2k, \dots, 2k \rangle}(\mathbf{Kc}_i) = k/(k+1).$$

5 See Williamson (2008a) for more on knowing when it is almost certain on one's evidence that one doesn't know (and iterations thereof), and Williamson (2008b) for some related semantic issues. The phenomenon discussed in the text involves the apparent loss of only one iteration of knowledge between premises and conclusion. However, the apparent absence of a given number of iterations of knowledge can cause doubts about all lower numbers of iterations, by a domino effect, since lack of knowledge

that one has $n + 1$ iterations implies lack of warrant to assert that one has n iterations (Williamson 2005c: 233-4).

6 Lewis (1980, 1986: 87) (with trivial differences of notation); for refinements see Lewis (1994).

7 The postulated closure of knowledge under logical consequence seems to imply that knowledge without belief is a far more widespread phenomenon than Lewis's exemplification of it with the timid student suggests — unless the closure of belief under logical consequence is also postulated.

8 Otherwise one could put $\mathbf{e} = \mathbf{p}$ in the Principle and derive a contradiction whenever $x \neq 1$, for $C(\mathbf{p} \mid \mathbf{ch} \ \& \ \mathbf{p}) = 1$.

9 I have changed the wording in unimportant ways. Hawthorne and Lasonen-Aarnio intend knowing future contingents not *ipso facto* to count as being inadmissibly connected to the future. The cases the qualification is intended to exclude are those 'in which there are no time-travellers from the future, clairvoyance by backwards causation etc.'

10 LC* even lacks the restriction in Low Chance to subjects not inadmissibly connected to the future.

11 Sainsbury (1997, 2002: 117-18) and Peacocke (1999: 310) make this point.

12 A further relativization may be needed to levels of confidence of belief or near-belief, as discussed in the book (98-99). Since this relativization can be treated in formal parallel with the relativization to bases, it is omitted here.

13 Suppose that in case α I know that Mary is married, on the basis of seeing that she is wearing a ring (on the appropriate finger), while in a case α^* that could very easily have occurred instead of α I believe falsely that Mary is unmarried, on the basis of seeing that she is not wearing a ring — she hardly ever wears her wedding ring, but on this occasion forgot to take it off (this is a variant of an example in Sainsbury (1997: 2002: 114)). Is this a counterexample to SAFETY? No. To a first approximation, I really do know in α that Mary is married only if wearing a ring is a far more reliable indicator of being married than not wearing one is of being unmarried, in which case the bases are not sufficiently close.

14 In Williamson (2005b: 484-7) I show formally how to limit the divergence by ensuring that only high chance propositions are true in all close worlds, but I agree with Hawthorne and Lasonen-Aarnio that in many cases of interest that model would avoid sceptical consequences only if it used a very unnatural measure of closeness.

15 Thanks to audiences at the University of Texas at Austin and the University of Santiago de Compostela for comments on versions of this material.