In the course of his famous exchange with Isaac Levi, Richard Jeffrey makes explicit the physics analogy behind the coinage "probability kinematics."

> In Physics, Dynamics is a contrary of Kinematics as well as of Statics: it is the first contrariety that I had in mind when I called Chapter 11 of *The Logic of Decision*, 'Probability Kinematics'. Take a see-saw, with fulcrum 2/3 of the way toward your end. If you push your end down two feet, the other end will go up three. That is kinematics: You talk about the propagation of motions throughout a system in terms of such constraints as rigidity and manner of linkage... When you talk about forces--*causes* of accelerations--you are in the realm of dynamics.[1]

Jeffrey's point, clearly, is that his formula for updating on a change in the probability of uncertain evidence is merely kinematic; it tells about the propagation of motion but nothing about the "causes of accelerations," i.e. about where such changes in probability come from or about whether they are rational.

This, I think, is quite correct. But Jeffrey was apparently not satisfied with according this modest status to his formula. From the outset, he presented it as having dynamic relevance and, specifically, as counting somehow "against" the strong foundationalism of C. I. Lewis.

Jeffrey states this motivation for probability kinematics explicitly when he discusses the history of the idea:

> Probability kinematics was first introduced for an in-house philosophical purpose: to show how, in principle, all knowledge might be merely probable, in the face of a priori arguments to the contrary, e.g., those of C. I. Lewis..., who saw conditioning as the only reasonable way to modify judgmental probabilities by experience.... Using probability kinematics, I aimed to show...how the familiar language of objective statement needed no supplementation by what C. I. Lewis... called "the expressive use of language..."[2]

In other words, if probability kinematics is rational, we do not need (and we certainly need not restrict ourselves to) certain evidence.[3] Conversely, Jeffrey seems to have assumed that that if we

---

[1]Richard Jeffrey, "Dracula meets Wolfman: Acceptance vs. Partial Belief," in Marshall Swain, ed., *Induction, Acceptance, and Rational Belief* (Dordrecht: D. Reidel, 1970), p. 172.

[2]Richard Jeffrey, "Conditioning, Kinematics, and Exchangeability," in *Probability and the Art of Judgment*, pp. 135-6.

[3]See also Richard Jeffrey, "Probable Knowledge," in Sidney A. Luckenbach, ed., Probabilities, Problems, and Paradoxes (Encino, CA: Dickenson Publishing Co., Inc., 1972), p. 97 and *Subjective Probability: The Real Thing* (Cambridge: Cambridge University Press, 2004), pp. 59-60.

do have certain evidence, probability kinematics is unnecessary.[4] On this view, probability kinematics and Lewis-style foundations are in a competition where each would render the other unnecessary.

Both of these competitive propositions are false. The fact that Jeffrey's Rule is probabilistically legitimate does nothing to remove the sorts of considerations that motivated Lewis. And the presence, even in all cases, of certain evidence does not render Jeffrey Conditioning superfluous. I shall focus most of all on the first of these claims, and I shall do so by way of considering a puzzle for Jeffrey Conditioning--the difficulty in ascertaining when the relevant posterior probabilities are rigid.

JC can be used to update the probability of H when some evidence E has shifted from one uncertain probability to another, but only when the posterior probabilities $P(H|\pm E)$ are the same in the old and new distributions. But, as Judea Pearl mildly points out, "[T]his condition...is not easy to test."[5] Jeffrey himself implied that the posteriors are rigid when H is not one of the propositions "directly" affected by some "passage of experience,"[6] but Pearl shows that there can be situations where the posteriors are not rigid but where it does not seem correct to say that H is "directly" affected by the passage of experience.[7]

This difficulty might seem to be merely one of access and hence a practical rather than a theoretical problem. We are not logically omniscient, and many logical and evidential matters may be beyond our ken in various instances. A theory is not necessarily flawed just because it emphasizes the importance of a probabilistic relation which, in some concrete cases, we find difficult to see clearly.

But the trouble here lies not with JC itself (which is, given the relevant assumptions, derivable directly from the Theorem on Total Probability) but rather with the attempt to use JC to render strong foundations unnecessary. Jeffrey was adamant that the change in the probability of E was *simply caused* by the passage of experience and that the experience could not be cashed out in propositional terms nor its relation to E modeled in probabilistic terms.[8] This aspect of radical probabilism lies at the heart of the difficulty in stating clearly when the posteriors are

---

[4]Richard Jeffrey, *The Logic of Decision* (Chicago: University of Chicago Press, 1965), p. 165. (Page numbering is from the second edition.)

[5]Judea Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (San Mateo, CA: Morgan Kaufmann Publishers, Inc., 1988), p. 64.

[6]Jeffrey, *Logic of Decision*, p. 168.

[7]Ibid., pp. 66-67. In Pearl's example, H is affected by the new sensory evidence via a different indirect route, not via E. Pearl's point is that one might not think of this route and might incorrectly apply JC, concluding that the probability of H remained unchanged, because of the irrelevance to H of E, where E is itself directly affected by the new experience. (I am here changing Pearl's notation to my own.) Pearl does make a slight misstep in explaining this example when he apparently conflates the color a piece of cloth appears to be with the color it actually is.

[8]Jeffrey, "Probable Knowledge," pp. 99-100. Jeffrey, *Logic of Decision*, pp. 184-5.

rigid. For if the passage of experience bears a causal relationship to one's beliefs and not an evidential relation, then the change it induces in the distribution is a surd. It is neither predictable nor criticizable.

The unpredictability point has been pressed by Mary Hesse, who argues from a pragmatic perspective that a confirmation theory is of no use unless it is possible to answer questions about "the effect of possible future evidence upon the probability of hypotheses." But, she points out, "The conditional probability $p(h_s/e_j)$ bears no regular relation to the new $p_r(h_s)$, because the physical observation of $e_j$ has an unpredictable causal effect upon the whole distribution."[9]

In other words, if Jeffrey were willing for the new experience to bear an evidential relation to the rest of the distribution, this would leave open the possibility that the relation could be modeled in probabilistic terms and hence that the changes induced in the distribution, including the conditions for JC itself--the change in E and the rigidity of P(H|E)--could be predicted. But this possibility is ruled out by the purely causal and subjective nature of the change. Not surprisingly, dynamic considerations limit the relevance of kinematics.

In a similar vein, Isaac Levi argues at some length that, given personalism, JC has no normative force. Since a personalist holds that S could have chosen one of many coherent distributions at $t_1$ and can choose one of many coherent distributions at $t_2$, there is no reason to require him to update his probabilities from $t_1$ to $t_2$ by following Jeffrey's Rule, by asking himself (for example) which propositions are directly affected by some experience and which are not. S might just as well choose some new coherent distribution or other bearing no regular relation whatsoever to the former one.[10]

Levi goes so far as to argue that a "passage of experience" is unnecessary and that Jeffrey has no principled way to distinguish a probability shift induced by a relevant "passage of experience" from one induced by a change in blood chemistry. Neither is more rational than the other nor epistemically more normative. Since there is no genuine normativity in personalism to begin with, no dynamic norms as to what sorts of things ought to count as evidence and drive one's other probabilities, normativity cannot be created *ex nihilo* by the introduction of a kinematic rule for updating on a change in the probability of uncertain evidence.[11]

Levi does not explicitly mention the issue of the rigidity of the posteriors, but again, that difficulty is related to his criticism. For if we restrict ourselves to considering Jeffrey's own examples involving some "passage of experience" which may affect the distribution in various ways, we may fool ourselves into thinking that this is a normative and orderly process. And if this were so, we might very well be able often to tell, by the seat of our pants as it were, how the experience would affect the relevant posteriors and the other propositions in the distribution.

---

[9]Mary Hesse, The Structure of Scientific Inference (Berkeley: University of California Press, 1974), pp. 122-3.

[10]Isaac Levi, "Probability and Evidence," in Swain, ed., *Induction, Acceptance, and Rational Belief*, pp. 140-2.

[11]Isaac Levi, "Probability Kinematics," *BJPS* 18 (1967), pp. 200-5. I should stress that Levi himself is by no means a strong foundationalist. He explicitly rejects Lewis-style foundations in favor of fallible foundations to which one nonetheless gives probability 1. See pp. 206-9.

This impression, in turn, lends some appearance of credibility to the claim that JC removes the motive for Lewis's *evidential* foundations, as experience still seems (despite Jeffrey's own insistence to the contrary) to be playing a quasi-evidential role. When we keep clearly in mind that entirely arational factors--indeed, non-experiential events inaccessible to the subject himself--can also cause a Jeffrey shift, all bets are off. Anything whatsoever could happen to the posteriors. Thus, again, the problem is not with probability kinematics itself but with the attempt to wring blood from a stone--to get dynamic value out of a kinematic rule.

A solution to the problem of when the probability of E changes but the posteriors remain rigid is to be found, ironically, in the very foundationalism Jeffrey rejects. Judea Pearl points out that, if the "passage of experience" is treated as a piece of evidence *e* and if A is some hypothesis and $B_i$ some intermediate-valued proposition in the distribution, then the posterior $P(A|B_i)$ will be rigid from the old to the new distribution only when A and *e* are conditionally independent modulo $B_i$, i.e. when

$prob(A|B_i, e) = prob(A|B_i)$.[12]

This relationship is sometimes described by saying that $B_i$ screens off *e* from A.

It can be shown, further, that if the only change in evidence is the addition of the new observational evidence, then this screening off relation is sufficient as well as necessary for rigidity.[13] The argument is as follows:

Take the definition of screening off according to which E SO p from H on some background k iff P(H|E & k) = P(H|E & k & p).

Suppose that k- is some body of given background evidence that does not include p. Now assume that, in the old evidence situation, p is not present. k- consists of S's given background information but does not include p. Then,

prob (H|E) =*def.* prob (H|E & k-)

(Note that this does not mean that E SO k- from E, as the k- is merely suppressed on the left side and expressed on the right.)

Assume that the only difference between prob and PROB is the addition of p at probability 1. Hence

PROB (H|E) = prob (H|E & k- & p)

Suppose that the addition of p does change the probability of E, so that

prob (E) $\neq$ PROB (E),

although neither value is 1 or 0. Suppose that E and H are intermediate-valued propositions in both probability distributions and that we are wondering whether we can use JC to model the

---

[12] Pearl, *Probabilistic Reasoning in Intelligent Systems*, p. 64.

[13] I developed this argument before I was aware of Pearl's work on the subject.

change (if any) in probability of H after the shift in probability of E.

Now, on the one hand, suppose that E SO p from H on k-. Then,

prob (H|E & k-) = prob (H|E & k- & p)

Therefore, by definition of SO, rigidity holds for E, i.e. prob (H|E) = PROB (H|E), since the two terms in the above statement equal, respectively, prob (H|E) and PROB (H|E). So, under the circumstances as described, E SO p from H is a sufficient condition for the rigidity of the posterior P (H|E).

Under these circumstances, it is also a necessary condition. For suppose, on the other hand, that E does not SO p from H on k-. Then,

prob (H|E & k-) ≠ prob (H|E & k- & p).

Hence,

prob (H|E) ≠ PROB (H|E).[14]

This argument expands upon Pearl's in several ways. Most obviously, it shows at greater length why conditional independence works in the way that Pearl claims it does. Furthermore, it expands his "only when" by pointing out conditions--in essence, an *evidentially expressed* set of conditions very much like those Jeffrey had in mind--in which the SO condition is sufficient as well as necessary for rigidity. And finally, it makes it particularly easy to explain the epistemic relevance of this result.

For strong foundationalists will admit as rational a change from one rational distribution to another only if there has been some change in the foundations. Intermediate-valued propositions, to a strong foundationalist, cannot rationally change their credibility merely because of some surd *cause*. Their probabilities are themselves derivative from their evidential relations to the given evidence. Since two perfectly rational subjects at any given time with the very same foundational evidence will have the same credibilities for all non-foundational propositions, they will only move from this static state to another if the foundations change. For to allow a change in propositions of intermediate probability without a change in the foundations would be, implicitly, to allow that two people with the same foundations can have different rational credibilities for non-foundational propositions.

It follows that a screening-off condition involving changes in foundational evidence provides a rule for the strong foundationalist as to the applicability of JC.[15] Suppose that E

---

[14]For JC to apply, P (H|-E) must also be rigid from prob to PROB, so the SO condition would have to apply with regard to -E as well.

[15]To express the rule in its entirety would, in my opinion, require also a discussion of changes that come about by deleting propositions from the foundations, since what I have direct access to at one time may be only indirectly accessible to me at a later time. It is easy to display deletion probabilistically in the above argument; it involves merely reversing the description of the probability of H given E in the distributions prob and PROB and proceding *mutatis mutandis*. It is more difficult to explain exactly what is meant by the deletion of a piece of evidence from

changes its probability rationally for S from one intermediate value to another. The strong foundationalist will look for a change in S's given evidence pertinent to E to justify this change. If the change is the addition of p, then the posteriors prob (H|±E) will remain rigid iff ±E SO p from H. But no such rule will be generally applicable either for moderate foundationalists or for personalists, both of whom allow propositions with probabilities of less than 1 to be evidentially basic. E could, then, change from one intermediate probability to another without a change in anything more foundational; hence, if we reject strong foundationalism, we must allow that there might be no evidential belief p that could be examined in its relation to E and H to see if the SO condition held.

Several objections will immediately spring to mind for those who do reject strong foundationalism. I propose to set aside here the one on which the most ink has been spilled in the literature already--whether in fact there is such a given element to experience and whether it can or should be treated evidentially. Obviously, if strong foundations are strictly unavailable (as Jeffrey himself clearly believed they often were), there is no point in discussing what the probabilistic situation would be if only we had them.

Two other objections can be discussed without rehearsing the entire debate between strong foundationalists and their critics. First, one might understandably wonder whether the screening off relation discussed here is any more accessible to the subject--even to a subject who is himself an epistemologist--than the rigidity of the posteriors themselves. May we not have solved the puzzle of when the posteriors are rigid only to introduce another, equally difficult of solution, as to when screening-off holds?

In fact, I think it fairly clear that the screening off relation is at least as psychologically accessible as the rigidity condition and that when we can tell easily and (seemingly) directly that rigidity holds, we are often implicitly accessing screening off. Jeffrey's own attempt to keep the discussion to situations where a "passage of experience" occurs and where some propositions are affected by it "directly" supports this contention.[16] So natural is it to link rigidity with screening off of the sort I have just discussed that James Hawthorne implies at one point that this sort of screening off just is the same thing as the rigidity condition.[17] But this is true only if, unlike Jeffrey himself, we allow the experience in question to be given evidential import.

But beyond this psychological fact of the matter, I am willing to say that it is simply epistemically better if our degrees of confidence are explicable in a principled fashion, and that this is true regardless of whether we access the underlying basis of those probabilities more or

---

the foundations, though I do not think such an explication is by any means impossible.

[16]In line with Pearl's criticism, it is important to stress that the new experience does not affect all the altered posteriors directly. Relatedly, and contra Jeffrey himself (*Logic of Decision*, pp. 168-9), JC can be used to model the effect on H of a change in E even if there are other propositions in the distribution for which the posterior probabilities on E do not remain rigid. JC need not be used to propagate the shift from E throughout the entire distribution. I am grateful to James Hawthorne for stressing this point in correspondence, for explaining in detail why it is right, and for drawing my attention to Jeffrey's own assumption to the contrary.

[17]James Hawthorne, "Three models of Sequential Belief Updating on Uncertain Evidence," *Journal of Philosophical Logic* 33 (2004), p. 93.

less easily than the intermediate probabilities themselves. And, as Levi points out, it is only by having some principled stance regarding rationality at a given time that we can take any rule for updating to be normative. By thinking in terms of foundational evidence and the ways in which it affects non-foundational propositions--sometimes directly, sometimes indirectly, sometimes by this or that route--we are enabled to think of both synchronic distributions and diachronic changes in a normative fashion, and this is all to the good.

What then, becomes of Jeffrey Conditioning? Here we return to the second competitive assumption mentioned at the outset--that the presence of certain foundations makes JC superfluous. Jeffrey always introduces JC after saying that given evidence is not always available, and Pearl explicitly conjectures that we need not bother with JC if we have certain evidence or can at least legitimately model our experience as certain evidence.[18]

This opposition between JC and foundationalism goes hand in hand with the assumption that, if we have certain foundations, other propositions will be based on them directly. At one time Jeffrey apparently believed that foundationalists require everything else to be based directly on the foundations. He implies that C. I. Lewis was motivated in his search for certain evidence by "an inability to see how uncertain evidence can be used," and he follows with an interesting misquotation of Lewis. Jeffrey quotes Lewis as saying, "If anything is to be probable, then something must be certain. The data which themselves support a genuine probability, must themselves be certainties."[19] But Lewis actually said, "The data which eventually support a genuine probability, must themselves be certainties."[20] Of course, Jeffrey also had other reasons for rejecting strong foundationalism, but the misquotation is interesting for the light it sheds on the assumed conflict between JC and strong foundationalism.

It certainly seems that Jeffrey Conditioning is useful in modeling the way in which a change in uncertain evidence changes other probabilities. We rarely think explicitly of our strictly certain, directly accessible foundational evidence. More often we are likely to ask ourselves how the probability of some higher-order hypothesis changes as the probability of some everyday, but still uncertain, proposition changes. If we must make a stark choice between conditioning directly on certain evidence and using Jeffrey Conditioning, it would be tempting to stick with uncertain evidence and to hope for the best as far as justifying intermediate probabilities by way of foundations with probability 1.

The response to this objection involves stating outright what has been more or less implicit heretofore: The epistemic value of Jeffrey Conditioning derives from the notion of

---

[18]Pearl, *Probabilistic Reasoning in Intelligent Systems*, p. 70.

[19]Jeffrey, "Probable Knowledge," p. 97. Jeffrey gives this same misquotation of Lewis repeatedly, e.g. *The Logic of Decision*, p. 167 and *Subjective Probability*, p. 60. In *Subjective Probability* he also quotes from Lewis's *Mind and the World Order*, where Lewis clearly rejects a requirement that everything be based directly on the foundations, and there Jeffrey does not appear to have had a misconception about foundationalist requirements on this point.

[20]C. I. Lewis, An Analysis of Knowledge and Valuation (LaSalle: Open Court, 1946), p. 186. The misquotation is pointed out by Timothy McGrew, *The Foundations of Knowledge* (Lanham, MA: Littlefield Adams, 1995), p. 71. McGrew discusses at more length there the charge that foundationalists are unable to assimilate uncertain evidence.

epistemic routing, and the contemplation of screening off helps to make epistemic routing explicit. The whole idea is that new evidence is relevant to the hypothesis H "through" or "by way of" its relevance to E.[21] If I have an experience as if of a tiger ramping upon me, this is relevant to the hypothesis that I will not live to see my next birthday by way of the (uncertain) proposition that there is indeed a tiger ramping upon me. In Judea Pearl's example, I have a particular color experience (a la Jeffrey) upon observing a piece of cloth by candlelight. Pearl notes that this experience is relevant both to the proposition that the cloth is indeed that color and also to the proposition that the candle wax is "a notoriously cheap brand known to produce flames deficient in violent content." As Pearl points out, the proposition about the brand of wax does not seem to be affected "directly" by the experience. But the color experience does bear upon the question of what light spectrum is being emitted by the flame and thereby upon the proposition about the brand of wax.[22]

Such routing facts are everywhere in a rational subject's epistemic economy, and they are both objectively real and epistemically crucial. We do not go directly from qualia to quantum theory. In fact, if all one had were an isolated experience like that of the ramping tiger and no additional relevant empirical knowledge--as, for example, in the case of an infant who has no idea what a tiger is, what a tiger looks like, or what a tiger might do--the evidential connection between the experience and the proposition about life expectancy would not exist. The importance of vast amounts of background knowledge--grounded, to be sure, in a host of occurrent memory-like experiences, but bearing on uncertain empirical facts--makes it clear that the concept of routing is of overwhelming importance to the understanding of evidence.

This is why we cannot do without Jeffrey Conditioning. When there is a change in the foundations, or even when we hypothetically imagine a change in the foundations, the new evidential situation affects some propositions more directly than others and affects some by way of its effect on others. This fine structure of a rational evidential corpus must not be ignored.

The reasons for the indispensibility of foundations are as cogent today as they were when Lewis articulated them. Bayesians, like everyone else, need foundations. But if foundations are indispensible, so too is the complex set of evidential relations by which they influence higher-level propositions. And as the second presentation will show, an understanding of Jeffrey Conditioning can help us to follow an otherwise tangled evidential thread backwards through the propositions that are intermediate--both in the sense of having non-extremal probabilites and in the sense of standing between the foundations and other propositions--to its origin in the foundations.

---

[21]John Earman refers to evidence as bearing on one proposition "only through" another in "Bayes, Hume, Price, and Miracles," in Richard Swinburne, ed., *Bayes's Theorem*, p. 103.

[22]Pearl, pp. 66-7.

*The mutual support objection*

In her book *Evidence and Inquiry*, Susan Haack argues that "pure" foundationalism is flawed because, lacking loops in its account of evidential relations, it cannot accommodate the sort of mutual support one finds in the cross-hatching entries in a crossword puzzle.

> How reasonable one's confidence is that 4 across is correct depends, *inter alia*, on one's confidence that 2 down is correct . . . [H]ow reasonable one's confidence is that 2 down is correct in turn depends, *inter alia*, on how reasonable one's confidence is that 4 across is correct.[1]

Since the "interpenetration of beliefs" illustrated by a crossword puzzle need not involve a vicious circularity, Haack suggests that pure foundationalism must be abandoned.[2]

The suggestion that foundationalism is mistaken because it rules out loops in evidential relations is quite remarkable: it would amount to at least a limited vindication of circular reasoning. The challenge, as we see it, is to model the phenomenon of mutual support without violating the intuition that circular reasoning is objectionable. From our point of view it will be an added advantage if the analysis makes use of the resources of strong foundationalism, to which we are committed on other grounds.[3]

*A flawed attempt at a solution*

In an earlier attempt to analyze mutual support from a foundationalist perspective, we argued that a foundationalist should understand it diachronically.[4] According to this analysis, it is not the case that all evidential relations can be exploited simultaneously. To avoid loops of support, a rational agent must choose to prioritize some of his beliefs rather than others; beliefs that are not prioritized will not receive the maximum support possible given total evidence if some of this support comes from beliefs that they are in turn supporting. At a different time, however, and without the introduction of new evidence, beliefs previously treated as purely supportive could be

---

[1]*Evidence and Inquiry* (Oxford: Blackwell, 1993), p. 86.

[2]By a somewhat different route, John Post comes to a similar conclusion. See his contribution to David Weissman, ed., *Discourse on the Method and Meditations on First Philosophy* (New Haven: Yale University Press, 1996), pp. 236-71.

[3]See Timothy McGrew, *The Foundations of Knowledge* (Littlefield Adams, 1995).

[4]Timothy McGrew, "How Foundationalists Do Crossword Puzzles," *Philosophical Studies* 96 (1999): 333-50. See esp. pp. 344-6. Timothy and Lydia McGrew, "Foundationalism, Transitivity, and Confirmation," *Journal of Philosophical Research* 25 (2000): 47-65. See esp. pp. 60-61.

taken as conclusions on which maximal available support was brought to bear.

There is *something* right about this analysis. The fact that two beliefs are positively correlated means that either of them could be taken, under appropriate circumstances, to supply some reason for the other; and it obviously does not follow from this that we would be justified in believing any two propositions with a sufficiently strong positive correlation. But to say this amounts, upon examination, only to an admission of the phenomenon of positive correlation together with a reiteration of the ban on circular reasoning. It does not tell us how to model mutual support.

The diachronic approach, unfortunately, does not really solve the synchronic problem as to what a rational agent's degrees of confidence will be. On a very modest interpretation, this approach considers only what the subject's *conditional* probabilities will be when he is temporarily "setting aside" some of his evidence and considering the probability of some proposition conditional on the remaining evidence. But this does not tell us what his degrees of confidence should actually be on total evidence.

More radically, the diachronic approach and the concept of prioritizing suggest that a rational subject may at one time literally and even deliberately *ignore* some of his relevant evidence for some of the propositions in his distribution and, at a different time, take that evidence into account for those propositions but ignore some evidence pertinent to others. That is, if a subject happens to be aware of evidential relevance that goes from A to B and also from B to A, he must pretend at one time that he is not aware of the evidential connection in one direction, fixing his credibilities accordingly, and, if he re-prioritizes, pretend that he is not aware of the connection in the other direction and have, in consequence, a different set of credibilities. But this answers the synchronic question at too high a price, for it violates the requirement that the subject be perfectly rational at the synchronic moment. We are strongly committed to what we would call the objectivity constraint on rational belief:

> *Disagreements regarding the probability of any proposition are due either to differences in the relevant evidence available to the disagreeing parties or to specific inferential failures on the part of at least one of the disputants.*

But if two subjects with the very same evidence can be equally rational despite assigning different probabilities to their inferred beliefs in virtue of prioritizing their beliefs differently, this condition does not hold.

Hence, the diachronic attempt to solve the question of mutual support either evades the question of actual synchronic credibilities on total evidence or deteriorates into subjectivism. The problem of mutual support, then, requires a different solution.

*Response to the mutual support objection*

To answer the mutual support objection in a satisfactory way, the foundationalist needs to construct a model in which the impact of foundational evidence on higher level beliefs in mutual support scenarios is displayed without generating circularity or violating the objectivity constraint. The most difficult part of this, it turns out, is the clarification of the notions of *support* and *circularity*. Both concepts can be clarified by reference to Diagrams 1 and 2.

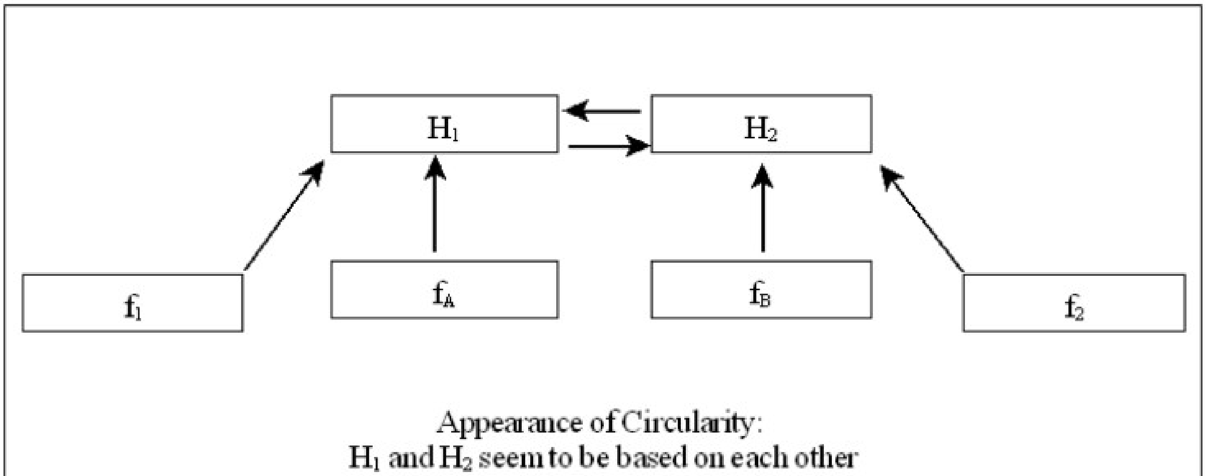


Diagram 1

In Diagram 1, four unproblematic foundational beliefs provide support for $H_1$ and $H_2$. The beliefs $f_1$ and $f_A$ support $H_1$ directly: their impact on $H_1$, both individually and jointly, is positive independent of whether $H_2$ is true. The higher level beliefs $H_1$ and $H_2$ are mutually relevant, but $\pm H_1$ screens off $f_1$ and $f_A$ from $H_2$. The situation is symmetrical with respect to $f_B$, $f_1$, and $H_2$; they provide, individually and jointly, direct positive support for $H_2$, but $\pm H_2$ screens off their impact on $H_1$.

The arrows in Diagram 1 do indeed give an impression that there is something circular going on. The eye travels from $f_1$ up through $H_1$ and across to $H_2$, but then doubling back we come back to $H_1$ again – as though, like the Baron Munchausen, $H_1$ could pull itself out of a probabilistic bog by its own hair. Clearly something is wrong with this picture. Yet we have stipulated that $H_1$ and $H_2$ are mutually relevant, so the arrow going from left to right and the arrow from right to left seem equally legitimate. (After all, when $P(H_1|H_2) > P(H_1)$, $P(H_2|H_1) > P(H_2)$.) How can we dispel the appearance of circularity here?

The central idea of the solution is that simple conditionalizing on the foundational beliefs propagates through the higher levels of the distribution in what are, essentially, Jeffrey conditionalizations. By attending to the screening condition, we are in a position to refine the notion of support: instead of saying that $H_2$ supports $H_1$ *simpliciter*, we can say more accurately that $H_2$ is a conduit through which $f_2$ and $f_B$ support $H_1$, and similarly that $H_1$ is a conduit by which $f_1$ and $f_A$ support $H_2$. The foundationalist's restriction on circularity is thereby similarly refined: no proposition may support itself, in the sense that no line of support may originate or pass through a belief node and also return to it.

The following distribution respects these constraints and illustrates the resolution of the problem of mutual support. Let $P(H_2) = .5$, and:

| $H_2$ | $P(H_1)$ |
|---|---|
| T | .8 |
| F | .4 |

| $H_2$ | $P(f_B)$ |
|---|---|
| T | .9 |
| F | .25 |

| $H_2$ | $P(f_2)$ |
|---|---|
| T | .6 |
| F | .4 |

| $H_1$ | $P(f_A)$ |
|---|---|
| T | .9 |
| F | .25 |

| $H_1$ | $P(f_1)$ |
|---|---|
| T | .6 |
| F | .15 |

Given this distribution, we can fix two of the foundational beliefs – for simplicity, $f_A$ and $f_B$ – and then watch what happens as we acquire $f_1$ and $f_2$. First, $P(H_1|f_A \& f_B) = .899$ (rounding here and elsewhere to three decimal places), and $P(H_2|f_A \& f_B) = .845$. Bringing in first $f_1$ and then $f_2$, we get $P(H_1|f_A \& f_B \& f_1) = .973$ , $P(H_2|f_A \& f_B \& f_1) = .869$ and then $P(H_1|f_A \& f_B \& f_1 \& f_2) = .976$, $P(H_2|f_A \& f_B \& f_1 \& f_2) = .979$. Since the shift is induced by the acquisition of certainties at the foundational level, order is immaterial: if we acquire $f_2$ first, we have $P(H_1|f_A \& f_B \& f_2) = .910$ and $P(H_2|f_A \& f_B \& f_2) = .891$, but when we then fold in $f_1$ we arrive at the same final numbers as before.

In what sense is either of these a Jeffrey shift? Due to the screening off conditions, a change in $H_1$ from one intermediate value to another is propagated to $H_2$ via a Jeffrey shift, and *vice versa*. The *ultimate* source of the change in the probabilities is the acquisition of foundational information, but in the intermediate propositions we still need Jeffrey conditioning to show how probabilities can coherently be redistributed.

The upshot is that we can show, by means of two different diachronic paths to a single final coherent distribution, how each of the propositions $H_1$ and $H_2$ can channel epistemic support to the other without any incoherence, circularity, or violation of the objectivity constraint. This solves the problem of mutual support.

All of this is well illustrated in Diagram 2, where the black arrows of the first diagram have been replaced by color-coded arrows that give us more insight into the underlying support relations. Critically, there is nowhere a cyclical path of the same color. This permits a visual representation not only of the positive correlation of $H_1$ and $H_2$, which is of course a symmetric relation and gives us the "mutual" in mutual support, but also of the channeling of epistemic support from more fundamental premises through $H_1$ to $H_2$ and *vice versa*, without circularity.[5]

---

[5]Each of the propositions at the foundational level has a probability conditional on either $H_1$ or $H_2$, and this likelihood fluctuates as some of the other foundational beliefs are acquired with certainty. Nothing about this is incompatible with strong foundationalism; unless some experiential foundational beliefs were expected more strongly under certain conditions than under others, the foundations could do no epistemic work higher up in the system of one's beliefs. (Try setting all of the conditional probabilities of all of the foundational beliefs in our distribution to .5.)

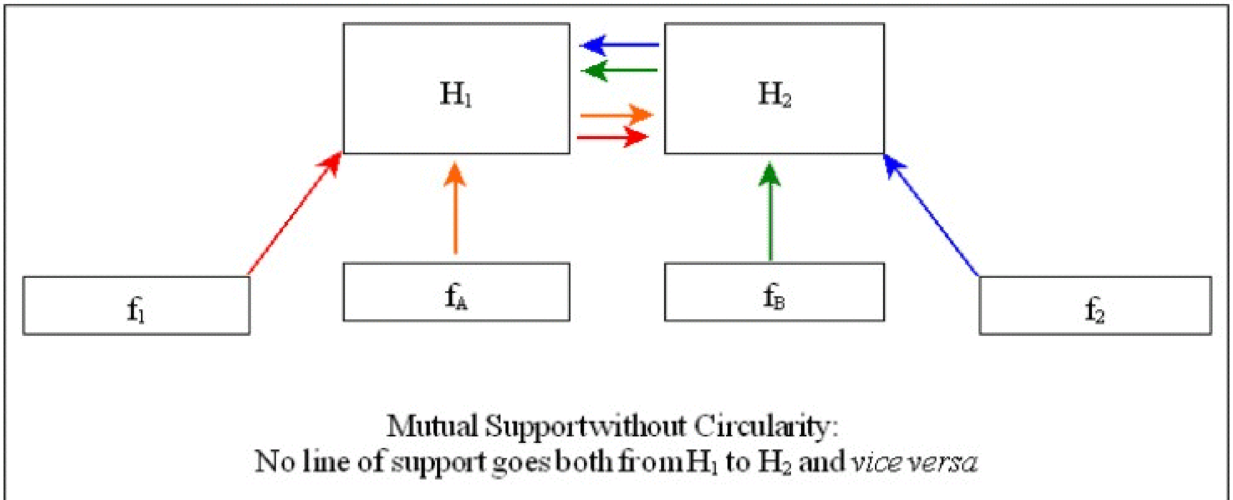


Diagram 2

*The basing relation*

Perhaps the most intriguing implications of this analysis of mutual support concern the concept of basing. If we conjoin an unqualified and unclarified statement that basing is transitive with an acknowledgment of synchronic mutual support, we will be forced to endorse loops of epistemic support, contrary to the norms of foundationalism. *Prima facie*, if A is based on B and B is based on A, then A is based on itself – the picture suggested by Diagram

What is needed, then, is a more detailed and sophisticated notion of basing. The foundationalist can and should insist that basing be thought of as taking place only within a line of support. On this conception, some propositions are based directly on the foundations while others are supported by way of intermediate-valued "conduits," propositions that channel the epistemic force of the foundations to other propositions. The line of support passes from the foundations, through conduit premises (if there are any) to some other proposition or propositions in the distribution. Understanding the epistemic underpinnings of the subject's rational credibility for a proposition is thus not simply a matter of listing all of the propositions that appear anywhere in any probabilistic relation either to it or to any of its premises. We must follow the lines of support.[6]

The notion of a conduit requires some explication. First, a conduit in some given subject's cognitive structure must be treated as such by the subject. No belief, foundational or non-foundational, can be a premise for a particular subject merely as a result of probabilistic facts, if he is not making use of those probabilistic facts inferentially. He must be basing B on A for A to be a premise for B, and it may well be that this (in one sense) psychological aspect of basing is *sui generis*, not admitting either of a purely causal or of a purely probabilistic analysis. And we are not inclined to capture this idea of basing by saying that the subject actually knows or must know metalevel facts about evidential and probabilistic relations, though we would say that a

---

[6]More generally, we should call these "lines of evidence," since evidence can be either positively or negatively relevant.

logically omniscient "upgraded version" of the subject would indeed know these facts.[7]

The more tractable aspect of what it means for a premise to be a conduit is the evidential and probabilistic one. Here we are indebted to the notion of a basis articulated by Hawthorne:

> An evidence basis is a partition – i.e., a mutually exclusive and exhaustive set of sentences. The evidential impact of an experience or observation *e* is supposed to be completely captured by the influence it has on the belief strengths of sentences of its evidence basis....The evidence basis captures the evidential import of *e* through its ability to "screen off" the rest of the agent's beliefs from direct influence by *e*.[8]

We would add, what is entirely compatible with Hawthorne's explication of a basis, that the propositions of the basis may have to be complex, consisting of the negations and assertions of more simple component propositions in all their permutations, such as {(A & B), (A & ~B), (~A & B), (~A & ~B)}. In that case, we wish to use the term 'conduit' for the simpler components as well as for the actual propositions of the basis as Hawthorne defines it. Moreover, we should note that the conduit propositions in such a basis must be evidentially relevant to the propositions to which they channel evidence, otherwise the concept of a basis consisting of conduit propositions can be trivialized.[9] Finally, as Hawthorne has emphasized, the basis need not screen off *e* from *all* other propositions in the distribution.[10] This means that to say that some partition is a basis for *e* is elliptical: A basis should be understood relative both to *e*, on the one hand, and to some particular other proposition or propositions in the distribution, on the other, even if it does not screen *e* from every other proposition.

Once we understand that premises can be conduits of evidence, it becomes clear that the actual probability of $H_1$ at t, conditional on *all* evidence pertinent to $H_1$, is not the relevant point to consider in understanding how or whether it is supporting $H_2$.[11] It is not as though it can be supportive only if its probability meets some k criterion. For it could be channeling positively relevant foundational evidence to $H_2$ even if its own probability is low. And some of the evidence pertinent to and supporting $H_1$ may be channeled by $H_2$ itself as a conduit. If this is the case, it is

---

[7]We discuss this issue at greater length in our forthcoming book *Internalism and Epistemology* (Routledge, 2006).

[8]James Hawthorne, "Three Models of Sequential Belief Updating on Uncertain Evidence," *Journal of Philosophical Logic* 33 (2004), p. 93.

[9]This condition blocks the swelling of a basis by the addition of irrelevant sub-components. If, in the context, C is quite irrelevant to H, then there is no point in tacking C or its negation onto the end of the propositions that constitute the basis, even if the resulting partition would also screen H from *e*.

[10]Hawthorne has emphasized this point in personal correspondence.

[11]It will, of course, be *pertinent* to $P(H_2)$, but only in the sense that we can calculate $P(H_2)$ in terms of $P(H_1)$ and the relevant conditional probabilities. But the question of *how* $H_1$ supports $H_2$ is different and more interesting.

doubly important that we not think of the probability of $H_1$ on total evidence as critical to its being or not being a premise for $H_2$. Nor will we come to understand the role $H_1$ plays with respect to $H_2$ in one's cognitive economy better simply by calculating $P(H_2)$ in terms of a partition $\{H_1, \sim H_1\}$. Any two propositions in the distribution can be "related" by such a calculation. This does not tell us whether or how either of them is acting as a premise for the other.

Strong foundationalists should say that any proposition A that is a premise for B and that is not itself foundational is a conduit of foundational evidence to B. This means that the notion of a conduit is not arcane to a foundationalist; he has been tacitly using it all along, even if he has not realized its importance for mutual support. In terms of Diagram 2, $H_1$ and $H_2$ are conduits of support to each other, and it is for exactly this reason that there is no circularity. Each supports the other only in the sense that and insofar as it channels evidential support to the other *unidirectionally* from elsewhere in the distribution and, ultimately, from the foundations. $H_1$ is based on $H_2$ only within the lines of support that run from $f_B$ and $f_2$ upwards, and $H_2$ is based on $H_1$ only within the lines of support that run from $f_A$ and $f_1$ upwards.

This notion – that basing occurs only within lines of evidence – clarifies the issue of transitivity. Basing is transitive, but that is because it occurs only within evidential lines and must always be understood relative to such a line. And lines of support are themselves acyclic. Within the line of support from $f_B$ and $f_2$, $H_1$ is based on $H_2$, $H_2$ is based on $f_B$ and $f_2$, and $H_1$ is based on $f_B$ and $f_2$ via their impact on $H_2$ as a conduit. There is here neither any loop nor any counterexample to transitivity. And the same is true *mutatis mutandis* for the lines that run from $f_1$ and $f_A$ to $H_2$ via $H_1$.[12]

Here, however, it may seem that we have smuggled in the assumption of foundationalism by stipulating that lines of support are acyclic. Why could they not contain loops, the anti-foundationalist might ask?

First, we should stress that the challenge was for a foundationalist to model mutual support without violating his own norms, so it is not relevant to point out that, in so doing, we have kept in place a ban on loops of support. If loops are not needed for modeling mutual support, anti-foundationalists will have to find some other reasons for questioning their exclusion.

Moreover, once we are doing the probability theory in detail, it becomes particularly difficult

---

[12]We want to distinguish the notion of basing from the notion of confirmation. Basing is transitive *simpliciter*, and this holds independent of whether sufficient conditions hold for confirmation to be transitive. Tomoji Shogenji has proven that, when ±Y screens off X from Z, and the inequalities $P(Y|X) > P(Y)$ and $P(Z|Y) > P(Z)$ both hold, then $P(Z|X) > P(Z)$ ("A Condition for Transitivity in Probabilistic Support," *BJPS* 54 (2003): 613-16). In the case where a given partition acts as a basis for *e* with respect to H, the items in the partition that take the place of ±Y might be complex, and although there are several useful extensions of Shogenji's theorem for such cases, they all require some additional constraints for the transitivity of confirmation. When the relevant extensions of Shogenji's screening condition are met, the transitivity of confirmation and the transitivity of basing will be co-exemplified. But Z may be based on $Y_1$, which is itself merely a component of sentences of a basis that screens X from Z; and if $Y_1$ is based on X, then Z is based on X, even when $Y_1$ does not by itself screen X from Z. And even the (possibly complex) sentences of the basis might be negatively relevant to Z, so 'basing' as we are conceiving it would not always require the transfer of support.

to see what could be meant by “loops of support.” To insist that there are real loops of support, and to insist that this means that we must jettison foundationalist norms, is either to speak inexactly or to create self-evident probabilistic incoherence. If all that is meant by a “loop” is that one might draw an arrow (representing basing) from A to B and an arrow from B to A, then we could well be talking about legitimate mutual support. But we have already shown that propositions can channel independent evidence to each other without violation of foundationalist norms, and that such channeling may be represented by arrows going in each direction, though they should be clearly separated into different evidential lines. Beyond this, there can be no benign meaning to the concept of anti-foundational loops of support.

Consider the situation in which $H_1$ is, on the background, positively relevant to $H_2$ and vice versa and in which the anti-foundationalist wishes to say that the two propositions support each other with a real loop, violating foundationalist requirements. Now suppose that the probability of $H_1$ goes up independent of $H_2$, that the relevant posteriors are rigid, and that the shift is propagated to $H_2$. This is all that can coherently be done, and it is entirely consistent with foundationalist norms, as would be a similar shift in the other direction as a result of new independent evidence for $H_2$. But if the notion of a real loop of support is to mean something more than this, it could only mean that we must allow a ratchet effect whereby, after a shift that originates in $H_1$ is propagated to $H_2$, the probability of $H_1$ goes up immediately *yet again* “because of” the rise in $H_2$, and so forth. Such a feedback loop creates incoherence, as it produces multiple, different probabilities for both propositions at a single point in time. We cannot have a stable, coherent probability for $H_1$ if $H_1$ is literally lifting itself by its bootstraps. Probabilistic support for $H_1$ must originate independently and have its impact only once.

We need to make it clear that no one is explicitly advocating this sort of *ad infinitum* self-support. In fact, Haack insists upon independent support for each of the two mutually supporting propositions.[13] But if real self-support, real bootstrapping, is not in view, then what is the alternative to foundationalism supposed to be? And where is the challenge in the phenomenon of mutual support? It appears that the whole problem can now be reduced to one of representation, to a mental image of arrows going from A to B and *vice versa*. But the strong foundationalist should never have conceded in the first place that support comes *from* uncertain evidence A to uncertain proposition B, as if the evidential force originated in A. He should, rather, have insisted all along that support goes *through* uncertain A to uncertain B. Once this is clearly understood, the confused image of loops of support can be resolved into a clearer one of unidirectional lines of support passing through two different propositions, and the mutual support objection to foundationalism collapses.

What all this goes to show is that Bayesian probability theory is something foundationalist epistemologists have needed all along. Evidence trees (or bushes) sprouting arrows in all directions are a poor substitute for a clear and distinct understanding of the probabilistic inner workings of the propagation of shifts in uncertain evidence. And epistemologists who know ahead of time just what they want to model (mutual epistemic support, for example) are in a particularly good position to exploit the resources of probability theory for their purposes.

If, then, the previous presentation showed that foundationalism is good for Bayesians, the moral here should be that Bayesianism,properly understood, is good for foundationalists.

---

[13]Haack, *Evidence and Inquiry*, p. 86.