Infinite Regresses of Justification

Oliver Black

INFINITE REGRESS ARGUMENTS are common in philosophy, but few philosophers have thought to say what an infinite regress argument is. I propose the following definition. Let "x_1" and "x_2" be unrestricted variables, "A" a schematic predicate, and "R" a schematic expression for a binary relation. Then an infinite regress argument is one that approximates to the following schema:

i) $(\forall x_1)\ [Ax_1 \rightarrow (\exists x_2)\ (Ax_2\ \&\ x_1Rx_2)]$ (premiss)

ii) $(\exists x_1)Ax_1$ (premiss)

iii) R is irreflexive (premiss)

iv) R is transitive (premiss)

v) $(\exists s)\ [\mathrm{Inf}(R(s))\ \&\ (\forall i)\ (i \in D(s) \rightarrow As_i\ \&\ As_{i+1}\ \&\ s_iRs_{i+1})]$

—where "s" and "i" are variables ranging respectively over sequences and the positive integers. (v) says that there is a sequence with infinite range, each of whose elements has the property A and stands in R to its successor. Note: the claim is that the sequence has infinite range, not just that it is infinite. For an infinite sequence may consist of an infinite iteration of a finite number of elements: infinite regress arguments do not concern repetitive sequences of that kind. (v) is derived in three stages: first, an inductive procedure is specified for generating from (i)–(iv) a sequence that satisfies the second conjunct of (v); second, the range of the sequence is proved to be infinite; third, (v) is inferred with the rule for introducing the existential quantifier. The deduction is straightforward.[1]

vi) $\sim$(v) (premiss)

vii) (v) & $\sim$(v) (&I)

viii) $\sim$(i) / $\sim$(ii) / $\sim$(iii) / $\sim$(iv) (RAA)

That is, the conclusion is the denial of one of the premisses (i)–(iv). *Reductio ad absurdum* also licenses an inference to the denial of (vi). But that conclusion

[1]It is given in Oliver Black, "Induction and Experience: an Alleged Infinite Regress," in *Mythes et Réalités de l'Activité Scientifique*, edd. G. Simon and J. Vienne (Lille: Univ. de Lille III, 1985). Reprinted in *Fundamenta Scientiae* 7, No. 3/4 (1987).

would never be drawn since it would make the last three steps of the argument redundant: ~(vi) is equivalent to (v).

Infinite regress arguments diverge from the schema to various degrees. They are seldom stated so formally, and even a rational reconstruction may differ in containing modal operators, in using conditional proof and *modus tollendo tollens* instead of &-introduction and *reductio ad absurdum,* or in other respects. The definition's vague word "approximates" is tended to leave room for these variations. Usually an infinite regress argument will conclude to the denial either of its premiss corresponding to (i) or of that corresponding to (ii). The premisses corresponding to (iii) and (iv)—those specifying properties of the relevant relation—are usually taken for granted and seldom made explicit.

The definition is, of course, intended to fit arguments that are generally agreed to be infinite regress arguments. To see that it does, consider Aquinas' First Way:

> It is certain as a matter of sense-observation that some things in this world are in motion. Now whatever is in motion, is moved by something else. . . . Moreover, this something else, if it too is in motion, must itself be moved by something else, and that in turn by yet another thing. But this cannot go on for ever. . . . And so we must reach a first mover which is not moved by anything: and this all men think of as God.[2]

Aquinas' reasoning can be represented as follows:

i_1) If a thing is in motion, it is moved by something that is in motion;

ii_1) Something is in motion;

iii_1) . . . is moved by . . . is irreflexive;

iv_1) . . . is moved by . . . is transitive;

v_1) There is a sequence, with infinite range, of things in motion each of which is moved by its successor;

vi_1) There is no such sequence;

vii_1) There both is and is not such a sequence;

$viii_1$) Not-(i_1); that is, there is an unmoved mover.

(i_1) is implied by the second and third sentences of the quoted passage. The second says that if a thing is in motion it is moved by something; and Aquinas clearly assumes that the mover is itself in motion, for otherwise the third sentence is a *non-sequitur*. (ii_1) embodies the first sentence, except for the reference to sense-observation. (iii_1) and (iv_1) are implicit in the argument, as is (v_1). (vi_1) corresponds to the claim that "this cannot go on for ever," (vii_1) is taken for granted and ($viii_1$) corresponds to the final sentence of the passage.

[2]St. Thomas Aquinas, *Summa Theologiae* (Rome: Leonine Edition, 1894), Ia.2.3. The translation is Anthony Kenny's, in *The Five Ways: St. Thomas Aquinas' Proofs of God's Existence* (London: Routledge & Kegan Paul, 1972).

Substitution in the schema yields an argument that articulates the traditional problem of infinite regresses of justification:

I) A belief is justified only if it stands in R_1 to a justified belief (premiss)

By "belief" I shall understand "token mental state of belief," and the statement that a belief B_1 stands in R_1 to belief B_2 may be read for the moment as "B_2 is a reason for B_1"; this reading will be refined later.

II) There are justified beliefs (premiss)

III) R_1 is irreflexive (premiss)

IV) R_1 is transitive (premiss)

V) There is a sequence, with infinite range, of justified beliefs each of which stands in R_1 to its successor (from (I)–(IV) by the specified procedure)

VI) There is no such sequence (premiss)

VII) There both is and is not such a sequence (&I)

VIII) $\sim$(I) / $\sim$(II) / $\sim$(III) / $\sim$(IV) (RAA)

In these terms the problem of infinite regresses of justification is this: which premiss of the argument is false and hence should be rejected at (VIII)? If (I) is false, a component of a form of foundationalism is true; it is true, that is, that there are beliefs that are foundational either in the strong sense that they are justified even if they do not have reasons, or in the weaker sense that they are justified even if they do not have reasons that are themselves beliefs, or in the still weaker sense that they are justified even if they do not have reasons that are beliefs that are in turn justified. These claims constitute only parts of foundationalist theories of justification; a full theory will also state that, and how, justified nonfoundational beliefs derive their justification from a relation to foundational beliefs.

Infinite regress arguments about justification are often taken to have a foundationalist conclusion. But the present argument admits other possibilities. If premiss (II) is false, a radical form of scepticism is true: there are no justified beliefs. (III) and (IV) together amount to the thesis that the reason-relation R_1 is a partial order. If either of these premisses is false, there can be circles of reasons; that is, there can be either (1) $X_1, \ldots, X_{n>1}$ such that, first, for $1 \leq i \leq n\text{-}1$, $X_iR_1X_{i+1}$ and, second, $X_nR_1X_1$, or (2)—the limiting case in which a circle collapses into a point—an X such that XR_1X. If in addition all the other premisses of the argument are true, a form of coherentism is true; it is true, that is, that a justified belief is justified by virtue of being one of a set of beliefs that form a circle of reasons. Also, if (III) is false, a component of another form of foundationalism is true: there are beliefs that are foundational in the sense of being reasons for themselves. This foundationalist thesis is thus compatible with the form of coherentism just mentioned.

All the premisses are intuitively appealing, but at least one must be false. The rest of the paper is a brief survey of the premisses with a view to deciding their truth-values:

Premiss (I). (I) raises a number of epistemological issues that cannot be settled here; I shall merely map the terrain and state my own views dogmatically.[3] The discussion is best broken into three questions:

Q1) Is a belief justified only if it has a reason?

Q2) Supposing a reason is required, must the reason itself be a belief?

Q3) Supposing there must be a reason that is a belief, must it in turn be justified?

(I) is true if and only if the answer to all three questions is yes. Since (I) is intuitively attractive, so is an affirmative answer in each case. The answer no therefore should be given to one of (Q1)–(Q3) only if grounds for it can be supplied that outweigh the plausibility of the answer yes.

The simplest way to establish the answer no to (Q1) would be to give an example of a belief that is justified without having a reason. Various recent writers seem to have thought that certain kinds of beliefs have instances that are immediately justified in this sense: beliefs about one's current mental states, beliefs about appearances, perceptual beliefs about physical objects, memory-beliefs, beliefs about truths of reason and general beliefs of common sense. But these examples are all controversial and hence in themselves provide at best weak grounds for the answer no. For the answer to be made plausible, grounds are needed in turn for thinking that certain beliefs of the relevant kinds are indeed justified immediately. One approach is the following. Many philosophers have thought that some beliefs, by virtue of having certain properties, give the believer some sort of privileged access to their subject-matter. The properties have often not been precisely distinguished; examples are incorrigibility, indubitability, directness, self-warrant and certainty. It might be maintained (i) that some beliefs of the kinds mentioned have at least one of these properties and (ii) that any belief with such a property is justified immediately. These two propositions together imply that the beliefs in question are justified without having a reason and hence that the answer to (Q1) is no.

There are two objections to this approach. First, for any belief that is supposed to be immediately justified, the claim that it has one of these properties is at least as uncertain as the claim that it is justified without having a reason. Hence (i) is too doubtful to support the contention that the proposed examples are indeed justified without having a reason. Second, even if it is allowed that they have some of the properties, it may be denied to follow that they are immediately

[3]They are argued in Oliver Black, "The Infinite Regress of Justification," PhD thesis, University of London, 1987, ch. 4.

justified; Williams, for instance, holds that what he calls intrinsic credibility—which seems to be the property of being justified without having a reason—is not implied by incorrigibility.[4] In that case the contention is unsupported for the reason that (ii) is false. The appeal to privileged access is therefore not a promising way to establish the answer no to (Q1).

A firmer basis for the answer is provided by the externalist conception of justification. Externalist theories need not deny that justification may involve reasons, but such a theory can be formulated to admit justification without reasons and thus to imply that the answer to (Q1) is no. Consider for example a causal version of externalism stating that a belief is justified if and only if it is appropriately caused. Whether this implies a negative answer to (Q1) depends on the explication of "appropriately caused"; it does so as long as the types of causation counted appropriate do not all involve the presence of a reason for the belief. The externalist conception of justification, however, establishes the answer no only if it is an acceptable position. My own view is that, whatever its merits as an account of knowledge, externalism fails as an account of justification since it implies that certain clearly irrational beliefs are justified.[5]

The appeals to the concept of privileged access and to externalism constitute the strongest cases for a negative answer to (Q1). Given that both are unsatisfactory, and given the intuition in favour of an affirmative answer, I maintain that the answer to (Q1) is yes.

An affirmative answer to (Q2), while likewise intuitively attractive, runs against ordinary usage. On entering my room I form the belief that someone has been in there since I left. You ask me why I think so and I answer, "There is a smell of cigarettes." This sentence seems to give a reason that may be sufficient to justify my belief, but it also says nothing about beliefs: it describes my *room*. But this fact of language is insufficient to undermine the intuition in favour of the answer yes to (Q2). The obvious way to establish the answer no is to point to something that, first, is not a belief, second, is a reason for a belief and, third, justifies the belief. Various candidates have been claimed to meet these conditions in certain circumstances: propositions, observations and perceptions, forms of awareness more primitive than belief, and—perhaps members of the last category—sense-experiences.

The thesis that some sense-experiences meet the three conditions is plausible enough, I believe, to outweigh the presumption favouring an affirmative answer to (Q2). Consider a man who has a visual experience of the kind he would have if he were seeing something white in optimal circumstances and who believes, on the basis of this experience, that he seems to see something white. It is natural to

[4]Michael Williams, *Groundless Belief* (Oxford: Blackwell, 1977), pp. 160–62.

[5]This objection is developed in Laurence Bonjour, "Externalist Theories of Empirical Knowledge," in *Midwest Studies in Philosophy, Volume V: Studies in Epistemology*, edd. P. A. French, T. E. Uehling Jr., and H. K. Wettstein (Minneapolis: Univ. of Minnesota Press, 1980), p. 62; Alvin Goldman, "What is Justified Belief?" in *Justification and Knowledge*, ed. G. S. Pappas (Dordrecht: D. Reidel, 1979), pp. 16–18; Mark Pastin, "Knowledge and Reliability: a Study of D. M. Armstrong's *Belief, Truth and Knowledge*," *Metaphilosophy* 9 (1978), 157; Frederick F. Schmitt, "Justification as Reliable Indication or Reliable Process?" *Philosophical Studies* 40 (1981), 416; and Ernest Sosa, "Nature Unmirrored, Epistemology Naturalized," *Synthese* 55 (1983), 57–58.

say, unless perhaps the situation is unusual in some way, that the experience constitutes a justifying reason for the belief.

The thesis is most likely to be attacked on the grounds that no experience is a reason for a belief and hence that the second condition is unfulfilled. One argument to this effect runs as follows: all reasons for beliefs have propositional content; no experience has propositional content; therefore no experience is a reason for a belief.[6] But this argument establishes its conclusion only if both its premisses are acceptable. Neither premiss is forceful enough to be accepted as it stands: each is acceptable only if it is itself established by argument. But I know of no convincing case for either.[7] In the absence of further objections to the thesis, and granted the preponderance of intuition in its favour, I maintain that the answer to (Q2) is no.

The obvious way to justify the answer no to (Q3) is to point to something that, first, is a belief, second, is a reason for a belief, third, justifies that belief and, fourth, is not itself justified. And the most promising way to seek to establish the claim that there are beliefs satisfying these four conditions is to appeal to a contextualist theory of justification. A good example of such a theory is the one proposed by Annis.[8] Roughly, Annis' view is that N's belief that-P is justified if and only if N is able to meet certain objections to the belief. Whether the belief is justified is relative to the question that is being raised about P, and this "issue context" determines the group of people whose objections must be answered. The objections raised by members of the objector-group are determined by the practices and norms of justification that obtain in their community.

The theory needs to be supplemented by an account of what it is to raise and to meet an objection to a belief. Roughly, to raise an objection is to state a belief that one holds and that one regards as incompatible with the belief to which the objection is made. Correspondingly, to meet an objection is to show either that it is false or that it is compatible with the belief in question. It is in the spirit of contextualism to explain showing as a matter of convincing the objector: N meets an objection to his belief just in case he convinces the objector that the objection is false or compatible with it.

With this addition Annis' theory can be used to construct an example that satisfies the four conditions sufficient for a negative answer to (Q3). Suppose that N holds two beliefs, B_1 and B_2, and that B_2 is a reason for B_1. Suppose there is just one relevant objection, O, to B_1. Suppose N is able to state B_2 in response to O. What conditions must be fulfilled by B_2 if by stating it N is to meet O? In particular, must B_2 be justified? Not according to the account just given of what it

[6]Variants of the argument are found in Laurence Bonjour, "Can Empirical Knowledge Have a Foundation?" *American Philosophical Quarterly* 15 (1978), 10; Nelson Goodman, "Sense and Certainty," in *Empirical Knowledge*, edd. R. Chisholm and R. Swartz (Englewood Cliffs, NJ: Prentice-Hall, 1973), p. 362; Joseph Margolis, "Skepticism, Foundationalism, and Pragmatism," *American Philosophical Quarterly* 14 (1977), 125–27; Ernest Sosa, "The Raft and the Pyramid," in *Midwest Studies in Philosophy, Volume V: Studies in Epistemology*, pp. 6–7; Michael Williams, *Groundless Belief*, pp. 28–29; and Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 1958), I.486.

[7]Indeed Edward Craig makes a good case against the second premiss in "Sensory Experience and the Foundations of Knowledge," *Synthese* 33 (1976), 12–13.

[8]David Annis, "A Contextualist Theory of Epistemic Justification," *American Philosophical Quarterly* 15, No. 3 (1978).

is to meet an objection. By this account, N will meet O by stating B_2 so long as the statement will convince the objector that O is either false or compatible with B_1; it is not necessary in addition that B_2 be justified, that is, according to Annis, that N be able to meet all relevant objections to B_2. Suppose the statement will convince the objector. Then N will meet O by stating B_2. It has been granted that N is able to state B_2 in response to O. Hence N is able to meet O. Since O is the only relevant objection to B_1, N is able to meet all relevant objections to B_1. Therefore, by Annis' view, B_1 is justified. Suppose B_2 is not justified. Then, simplifying slightly, B_2 is a belief, it is a reason for B_1, it justifies B_1 and it is not itself justified. (The simplification is in the third claim; it is closer to Annis' theory to say that what justifies B_1 is not B_2 but the facts that N is able to state B_2 in response to O and that the statement will convince the objector.)

Contextualism establishes the answer no to (Q3) only if it is an acceptable account of justification. But it is unacceptable because it fails to explain why a belief's being justified is conducive to its being true. The thesis of truth-conduciveness is best expanded thus: a belief is more likely to be true if it is justified than if it is not. The fatal flaw in contextualism is that it fails to explain why this proposition is true. Why should the fact that a believer can meet the relevant objections make his belief more probably true than it would be if he could not?[9] But contextualism provides the most promising case for the answer no to (Q3). So, granted the intuitive presumption in favour of an affirmative answer, I maintain that the answer to (Q3) is yes.

To summarize: Premiss (I) of the infinite regress argument is true just in case the answer to each of (Q1)–(Q3) is yes. I have claimed that (Q1) and (Q3) have affirmative answers but that, since some sense-experiences are justifying reasons for beliefs, the answer to (Q2) is no. In that case (I) is false and hence should be rejected in the final step of the argument.

Premiss (II). (II) says that there exist justified beliefs. This is so obviously true that it is acceptable without being established by argument.[10] Indeed it is plausible to think that any argument for (II) will start from assumptions less certain than (II). Conversely, the view is plausible that any sceptical attack on (II) will use assumptions more controversial than (II). In that case the right response to such an attack is to turn it on its head and argue, by *modus tollens,* that, since (II) is obviously true, one or more of the sceptic's assumptions is false. This holds in particular for a sceptical use of the present infinite regress argument to infer (II)'s denial.

To claim that (II) is obvious is not to be enslaved to ordinary language. It is compatible with recognizing that intuitive judgments may be abandoned, in the course of developing a theory, for the sake of satisfying methodological requirements such as consistency and simplicity. The present argument may itself be seen as a moment in a dialectic between intuitions and theoretical constraints leading to a coherent view of justification: the *reductio ad absurdum* forces a

[9]This objection is developed in Oliver Black, "Justification and Context," in *Logic, Philosophy of Science and Epistemology,* edd. P. Weingartner and G. Schurz (Vienna: Hölder-Pichler-Tempsky, 1987), pp. 325–27.

[10]There is an appearance of contradiction here with the affirmative answer given to (Q1). I deal with this point in Black, *The Infinite Regress of Justification,* pp. 113–16.

revision of intuitions in the interest of consistency. But this can be admitted without (II)'s being regarded as a serious candidate for rejection; there are other premisses to choose from.

Premiss (III). To evaluate premisses (III) and (IV) it is necessary to make more precise the interpretation of "R_1." "B_1 stands in R_1 to B_2" has so far been read "B_2 is a reason for B_1." But there are various kinds of reason connected with beliefs. I shall define R_1 as follows:

> DR_1) A belief B_1 stands in R_1 to belief B_2 if and only if there exist a person N and propositions[11] P and Q such that (a) B_1 is N's belief that-Q, (b) B_2 is N's belief that-P, (c) P confirms Q and (d) B_1 is based on B_2.

This definition has two consequences. First, R_1 is defined only for beliefs as arguments. Now (Q2) asked whether, supposing a reason is required for a belief to be justified, the reason must be a belief. I answered no on the grounds that some sense-experiences are justifying reasons for beliefs. But, given that "reason" is understood in the present sense, the answer turns out to be a trivial yes. This conflict is easily avoided by extending the definition to admit experiences into R_1's range, but it will be simpler in what follows to hold to the present version. And I shall assume from now on that . . . confirms . . . relates only propositions. Second, not only is R_1 defined solely for beliefs, the definition applies solely to beliefs held by the same person. On the present interpretation premiss (I), which says that a belief is justified only if it stands in R_1 to a justified belief, entails that the same person holds both beliefs. Likewise (V), (VI) and (VII) concern only sequences of beliefs held by a single believer.

While premiss (III), which says that R_1 is irreflexive, is intuitively attractive, it is not, unlike (II), so forceful that it can be granted without argument: (III) is to be accepted if and only if established by argument. (DR_1) suggests two arguments for (III), one appealing to (c), the clause about confirmation, the other to (d), the clause about basing. Both are *reductiones ad absurdum:*

The first argument is this. Suppose R_1 is not irreflexive. Then there is a belief B that bears R_1 to itself. Let P be the propositional object of B. By (c), P confirms itself. But confirmation is irreflexive. So P does not confirm itself. So P both does and does not confirm itself. Therefore R_1 is irreflexive.

This argument establishes (III) only if the thesis that confirmation is irreflexive is acceptable. Since the other steps of the argument are uncontroversial, this necessary condition is also sufficient: the argument establishes (III) if and only if the thesis is acceptable. While, like (III), the thesis has some intuitive appeal, it is also like (III) in being acceptable if and only if established by argument. The present argument therefore establishes (III) if and only if supplemented by an argument establishing the thesis about confirmation. I know of no persuasive argument for the thesis.

The second argument for (III) runs as follows. Suppose R_1 is not irreflexive. Then there is a belief B that bears R_1 to itself. By clause (d) of (DR_1), B is based

[11] I assume for simplicity that propositions are the objects of beliefs; the discussion can be recast to accord with theories of belief that deny the assumption.

on itself. But . . . is based on . . . is irreflexive. So B is not based on itself. So B both is and is not based on itself. So R_1 is irreflexive.

As with the argument from irreflexivity of confirmation, this establishes premiss (III) if and only if the thesis that basing is irreflexive is acceptable. It has more intuitive appeal than the corresponding thesis about confirmation, but again it is to be accepted if and only if established by argument. Hence, as before, the present argument establishes (III) if and only if reinforced by an argument establishing that basing is irreflexive. To provide one, it is necessary first to give some account of basing. I propose this rough definition:

DBa) N's belief that-Q is based on his belief that-P = df the fact that N believes Q is explained by the fact that there is an appropriate causal chain from his belief that-P to his belief that-Q,[12]

where, for any P and Q, "The fact that-P is explained by the fact that-Q" entails P and entails Q, and where "There is a causal chain from X to Y" means "X stands to Y in the proper ancestral of the relation . . . causes. . . ."

In the light of (DBa) an argument is available for the thesis that basing is irreflexive: Suppose basing is not irreflexive. Then by (DBa) there exist N and P such that the fact that N believes P is explained by the fact that there is an appropriate causal chain from his belief that-P to his belief that-P. Hence there is an appropriate causal chain from the belief to itself. But "there is a causal chain from . . . to . . ." expresses an irreflexive relation. So therefore does "there is an appropriate causal chain from . . . to. . . ." Hence there is no appropriate causal chain from the belief to itself. So there both is and is not an appropriate causal chain from the belief to itself. Hence basing is irreflexive.

This argument establishes the irreflexivity of basing only if the proposition is acceptable that "there is a causal chain from . . . to . . ." expresses an irreflexive relation. Let the other steps be granted; except for the appeal to (DBa), they are uncontroversial. Then it is also true that the argument establishes its conclusion if the proposition is acceptable. The proposition is intuitively so forceful that it may be accepted without itself being established by argument. It follows that the present argument establishes the thesis that basing is irreflexive, hence that the thesis is acceptable, hence that the argument from the thesis, reinforced by the present argument, establishes premiss (III), and hence that (III) is acceptable. This completes the case for (III).

While the proposition that "there is a causal chain from . . . to . . ." expresses an irreflexive relation is immediately acceptable, it is instructive to consider what

[12]Causal theories of basing are the most widely held; see David Armstrong, *Belief, Truth and Knowledge* (Cambridge: Cambridge Univ. Press, 1974), ch. 6, esp. pp. 79–85; Robert Audi, "Foundationalism, Epistemic Dependence, and Defeasibility," *Synthese* 55 (1983), 126; Robert Audi, "Psychological Foundationalism," *The Monist* 62 (1978), 596; R. A. Fumerton, "Inferential Justification and Empiricism," *The Journal of Philosophy* 73 (1976), 564; Donald McQueen, "Belief and Reasons for Belief," *Proceedings of the Aristotelian Society, Supplementary Volume* 47 (1973), 70; George S. Pappas, "Basing Relations," in *Justification and Knowledge*, pp. 56–59; John Pollock, *Knowledge and Justification* (Princeton: Princeton Univ. Press, 1974), p. 48; Ernest Sosa, "Epistemic Presupposition," in *Justification and Knowledge*, p. 88, Marshall Swain, *Reasons and Knowledge* (Ithaca: Cornell Univ. Press, 1981), ch. 3; and Bernard Williams, "Deciding to Believe," in his *Problems of the Self* (Cambridge: Cambridge Univ. Press, 1973), pp. 141–42.

grounds might be given for it. The obvious course is to argue from the premisses that . . . causes . . . is irreflexive and that . . . causes . . . is transitive: (1) "There is a causal chain from X to Y" means "X stands to Y in the proper ancestral of . . . causes. . . ." (2) Suppose, for arbitrary X, there is a causal chain from X to X. (3) . . . causes . . . is irreflexive. (4) . . . causes . . . is transitive. (5) By (1) and (2), X stands to X in the proper ancestral of . . . causes . . . . (6) By (4) and (5), X causes X. (7) By (3), X does not cause X. (8) By (6) and (7), X both does and does not cause X. (9) By RAA, there is no causal chain from X to X. (10) By the arbitrariness of X, "there is a causal chain from . . . to . . ." expresses an irreflexive relation.

As before, this argument establishes its conclusion if and only if premisses (3) and (4), which jointly state that . . . causes . . . is a partial order, are to be accepted. (3), like the argument's conclusion, is—*pace* those who believe that God is *causa sui*—immediately acceptable. So the case turns on (4). The thought that . . . causes . . . is transitive is intuitively attractive. Unlike (3), however, (4) is acceptable if and only if established by argument. But an adequate argument can be supplied only on the basis of a theory of causality, which cannot be attempted here. So far, then, (4) is not to be accepted. Hence the present argument does not establish—though, given (4)'s plausibility, it does reinforce—the proposition that "there is a causal chain from . . . to . . ." expresses an irreflexive relation. Nevertheless that proposition, as has already been said, is acceptable without being established by argument. Hence the fact that it has not been so established does not prevent it from being used to establish (III).

As regards (III), then, the result is this: R_1 is irreflexive because basing is; basing is irreflexive because nothing is linked by a causal chain to itself; and the reason for this may be that . . . causes . . . is a partial order.

Premiss (IV). I shall now show that (IV), which says that R_1 is transitive, is false. Of the premisses in the infinite regress argument, (IV) has the weakest intuitive appeal. All arguments to an infinite regress of justification involve an assumption of transitivity like (IV), but it is seldom explicit. Proponents of such arguments generally seem, not to think that their premiss of transitivity is plausible, but to be unaware that they are using it.

(IV) nevertheless has some intuitive plausibility, and it is worth considering its source. An attractive explanation why we are inclined to believe that R_1 is transitive is that we assimilate it to the relation . . . is physically supported by . . . and think that this is transitive. To the extent that the analogy is persuasive, its cogency seems due in part to the fact that R_1 is analysed in terms of the basing-relation. To talk of one belief's being based on another is to use metaphorically a relational expression whose primary application is to physical objects; and . . . is physically based on . . . is close in meaning to . . . is physically supported by. . . .

To say that R_1 is transitive is to say that, for any X, Y, Z, if XR_1Y and YR_1Z, XR_1Z. The form of the refutation of (IV) will depend on the interpretation of the conditional operator here. I shall interpret it modally, so that (IV) is equivalent to the proposition that it is impossible that there should be X, Y, Z such that XR_1Y, YR_1Z and not-(XR_1Z). Thus construed, the premiss is refuted by showing that it is possible for there to be such X, Y, Z.

One way of showing this is to produce an example, that is, to describe three jointly possible beliefs that satisfy the condition. But that by itself would be an unilluminating refutation: merely providing an example does not *explain* the failure of transitivity, and an explanation is necessary to allay the intuition that R_1 is transitive. An explanatory demonstration that R_1 is not transitive will be given if it is shown either that . . . confirms . . . or that . . . is based on . . . is nontransitive. For suppose . . . confirms . . . is nontransitive. By the chosen analysis of transitivity, it is possible that there are propositions P, Q, R such that P confirms Q, Q confirms R, and P does not confirm R. But, if this is possible, the situation is also possible in which, in addition, there is a person N who believes these propositions, whose belief that-R is based on his belief that-Q and whose belief that-Q is based on his belief that-P. In that situation, by (DR_1), N's belief that-R stands in R_1 to his belief that-Q and his belief that-Q stands in R_1 to his belief that-P. But his belief that-R does not stand in R_1 to his belief that-P, for these two beliefs do not meet condition (c) of (DR_1). Thus, if . . . confirms . . . is nontransitive, so is R_1. And the nontransitivity of confirmation explains that of R_1. Parallel reasoning applies to the nontransitivity of basing.

I shall refute (IV) by showing that confirmation is nontransitive. Consider these three propositions:

P1) Paul is a logician who has forgotten Zorn's lemma.

P2) Paul is a logician.

P3) Paul can state Zorn's lemma.

(P1) confirms (P2), which confirms (P3); but (P1) does not confirm (P3).[13] The explanation why transitivity fails here is that (P1) defeats (P2)'s confirmation of (P3), defeat being characterized as follows: for any propositions P, Q, R, to say that P defeats Q's confirmation of R is to say that, first, Q confirms R and, second, (P & Q) does not confirm R.

Another way of showing that . . . confirms . . . is nontransitive is to appeal to probabilistic confirmation. For any P, Q, there is a threshold value v such that P confirms Q if the probability of Q, given P, is greater than v. Pretend for simplicity that v is 0.5 for all pairs of propositions. Then:

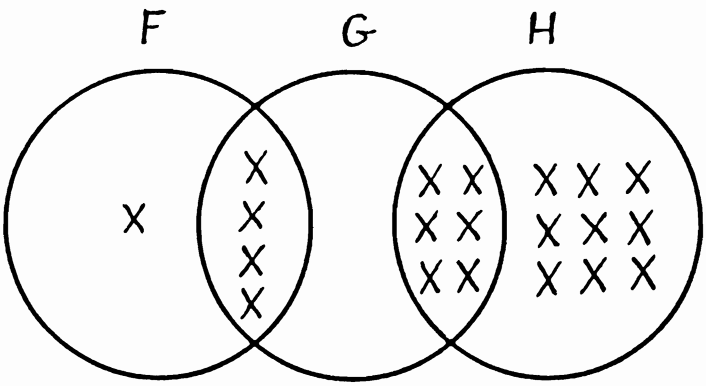
Pr) For any P, Q, P confirms Q if $pr(Q/P) > 0.5$.

Consider propositions (P4), (P5), (P6) such that, first, $pr((P5)/(P4)) = 0.8$, second, $pr((P6)/(P5)) = 0.6$, third, $pr((P4) \& (P6)) = 0$, and fourth, $pr((P4)) = 0.25$. Then, by (Pr), (P4) confirms (P5) and (P5) confirms (P6). But $pr((P6)/(P4)) = pr((P4) \& (P6))/pr((P4)) = 0/0.25 = 0$. (P4) and (P6) therefore fail to satisfy the condition given by (Pr) for one proposition to confirm another. (Pr) however states not a necessary but a sufficient condition for confirmation. So it may be that (P4) confirms (P6) by virtue of their meeting some other sufficient condition. But

[13]For a similar argument applied to justification, see Peter D. Klein, "Knowledge, Causality, and Defeasibility," *The Journal of Philosophy* 73 (1976), 806–807.

assume they meet no other condition sufficient for (P4) to confirm (P6). Then (P4) does not confirm (P6). It follows that confirmation is not transitive.[14]

This argument represents an actual state of affairs. Consider these sets:



Let X be a cross selected at random from the union of F, G and H. Now interpret (P4) as "X is in F," (P5) as "X is in G" and (P6) as "X is in H." These propositions exhibit the specified assignment of probabilities, and there is no way in which "X is in F" confirms "X is in H."

Both the argument from defeasibility and the one from probability provide a convincing case that confirmation is not transitive. It is worth noting two further arguments which reinforce this conclusion, although by themselves they are not decisive:

The first is similar in structure to the argument from probability and runs like this. The following principle is true:

E) For any P, Q, P confirms Q if Q explains P.

Consider propositions (P7), (P8), (P9) such that (P9) explains (P8), (P8) explains (P7) and (P9) does not explain (P7). By (E), (P7) confirms (P8) and (P8) confirms (P9). But (P7) and (P9) fail to satisfy the condition specified by (E) for one proposition to confirm another. (E), however, states only a sufficient condition for confirmation, so it may be that (P7) confirms (P9) by virtue of their satisfying some other sufficient condition. But assume they do not. Then (P7) does not confirm (P9). It follows that . . . confirms . . . is not transitive.[15]

To this argument, however, it might be objected that the state of affairs it represents is impossible. If (P9) explains (P8) and (P8) explains (P7), the objection goes, (P9) must explain (P7), for explanation is transitive. The question whether . . . explains . . . is a transitive relation cannot be pursued here. If the answer is

[14]For similar arguments, see Klein, p. 806; Keith Lehrer, "Justification, Explanation, and Induction," in *Induction, Acceptance, and Rational Belief,* ed. M. Swain (Dordrecht: D. Reidel, 1970), pp. 122–23; and John F. Post, "Infinite Regresses of Justification and of Explanation," *Philosophical Studies* 38 (1980), 39–40.

[15]Compare the use in Post, p. 40, of the concept of inference to the best explanation to cast doubt on the transitivity of justification.

yes, the objection stands and the argument falls. If the answer is no, the converse is the case, for there are no other plausible objections to the argument.

The arguments from probability and explanation appeal to principles—(Pr) and (E)—stating a sufficient condition for confirmation. The final argument for the nontransitivity of . . . confirms . . . uses one stating a necessary condition:

L) For any P, Q, P confirms Q only if P is relevant to Q.

Consider propositions (P10), (P11) and (P12) such that (P10) confirms (P11), (P11) confirms (P12) and (P10) is not relevant to (P12). By (L), (P10) does not confirm (P12). So confirmation is not transitive.

This argument also is exposed to the objection that it does not represent a possible situation. Since (P10) confirms (P11), it follows by (L) that (P10) is relevant to (P11). Likewise (P11) is relevant to (P12). But then, it may be said, (P10) must be relevant to (P12), for relevance is transitive. Again, I shall not pursue the question whether . . . is relevant to . . . is a transitive relation, so the position is the same as before: in the absence of other cogent objections, the argument stands if and only if relevance is nontransitive.

The upshot regarding (IV) is this. Of the four arguments, each of the first two, from defeasibility and probability, provides by itself a convincing case that . . . confirms . . . is not transitive and thus refutes (IV). The arguments from explanation and relevance reinforce the conclusion about confirmation, but whether they are to be accepted depends on the power of the objections that explanation and relevance are transitive. (IV) then is false and should therefore be rejected at step (VIII) of the infinite regress argument.

An argument with false premisses is a good argument only if it discharges them. The infinite regress argument has room to discharge only one premiss. But I have maintained that both premisses (I) and (IV) are false. If that is so, the argument as it stands is defective: if it concludes to the denial of (IV), it leaves (I) undischarged; if it concludes to the denial of (I), it leaves (IV) undischarged; and, if it concludes to the denial of some other premiss, it has two undischarged false premisses. Someone who wanted to use the argument to reach the negation of (IV) might seek to reconstrue the terms of the argument in such a way that (I) is true. But to employ the argument to refute (IV) would be unusual; as has already been said, in infinite regress arguments the premiss of transitivity is standardly taken for granted. I shall therefore consider only the corresponding reconstrual in the case of (IV). That is, I shall consider how the infinite regress argument may be reinterpreted to ensure that (IV) is true and thereby to allow the argument to conclude to the denial of one of the other premisses without violating, at least in the case of (IV), the rule that a false premiss must be discharged.

There are two possibilities; both involve a reconstrual of R_1. The first proposal is to subtract from the explication of R_1 that clause by virtue of which the nontransitivity of R_1, as presently conceived, follows from the nontransitivity of confirmation. At present R_1 is interpreted in terms of the definition (DR_1). The offending clause in (DR_1) is (c), P confirms Q. If (c) is removed, the resulting definition still characterizes a genuine concept of a reason. But, if R_1 is interpreted in terms of the new definition, which does not involve confirmation, the nontransitivity of . . . confirms . . . does not imply the falsity of premiss (IV). The second proposal is to reconstrue "R_1" as expressing R_A, the proper ancestral of the relation it now signifies. R_A, like all proper ancestrals, is transitive.

Premiss (VI). (VI) denies that there exists a sequence, with infinite range, of justified beliefs each of which stands in R_1 to its successor. For simplicity I shall call such a sequence a "J-sequence" and use "infinite sequence" to abbreviate "sequence with infinite range"; there will be no occasion in what follows to consider infinite sequences that do not have infinite range. I shall now argue that (VI) is true. As noted earlier, no one propounding the infinite regress argument would use it to reject (VI) in the conclusion, for then the last three steps of the argument would be redundant: since (VI) is the contradictory of (V), the conclusion would be identical to (V). But it follows from this, together with the fact that an argument is defective if it contains an undischarged false premiss, that the falsity of (VI) is sufficient for the infinite regress argument to be a bad one. For suppose (VI) is both false and rejected in the conclusion. Then the argument is bad by virtue of containing redundant steps. Suppose, on the other hand, that (VI) is false and not rejected in the conclusion. Then the argument is bad by virtue of containing a false premiss that is undischarged. To evaluate the infinite regress argument, therefore, it is essential to decide (VI)'s truth-value.

(V), which is derived from premisses (I), (II), (III) and (IV), entails:

Va) There exists an infinite sequence of justified beliefs held by the same person.

That they are held by the same person follows, as has already been said, from (DR_1). (Va) in turn entails:

Vb) There exists an infinite sequence of beliefs held by the same person.

Conversely (VI) is entailed by ~(Va), which is entailed by ~(Vb). It is convenient to discuss (VI) by first considering (Vb), then considering that part of (Va) that exceeds (Vb) and finally considering that part of (V) that exceeds (Va). Thus the first question is:

QVb) Does there exist an infinite sequence of beliefs held by the same person?

The second question is:

QVa) Supposing that such sequences exist, does any of them comprise only justified beliefs?

And the third question is:

QV) Supposing that there exist infinite sequences of justified beliefs held by the same person, is any of them a J-sequence?

This division of the discussion is parallel to that used earlier to address premiss (I). (VI) is true if and only if the answer to any of (QVb)–(QV) is no. I shall maintain that, while the answer to (QVb) and (QVa) is yes, (QV) has a negative answer.

The answer to (QVb) is yes, so long as "beliefs" is understood to embrace merely dispositional as well as occurrent beliefs. This claim has two components:

A) There exists an infinite sequence of beliefs all of which are held by the same person and some of which are merely dispositional;

B) There does not exist an infinite sequence of occurrent beliefs held by the same person.

(A) and (B) together entail that any sequence of the kind specified by (A) contains an infinite number of merely dispositional elements.

(A) is established by an example. Let S^P be the infinite sequence of propositions "2 is greater than 1," "3 is greater than 1," and so on. I believe each of these propositions. Hence corresponding to S^P there is an infinite sequence S^B such that each S^B_i is a belief of mine whose object is S^P_i. Clearly there are some elements of S^P that I never consciously consider; the elements of S^B corresponding to these are merely dispositional beliefs. S^B therefore is an infinite sequence of beliefs all of which are held by the same person and some of which are merely dispositional.[16]

(B) is an obvious empirical truth. We find, from introspection and observation of others, that every human mind is in only a finite number of conscious states throughout its existence and hence that no one ever holds more than a finite number of occurrent beliefs. There are of course difficulties in counting mental states and, sometimes, in determining whether a state is conscious or not; consciousness is plausibly regarded as a matter of degree. But neither of these areas of obscurity casts doubt on (B).

S^B also establishes the answer yes to (QVa). It is necessary here to distinguish between an actional and a statal sense of "justified": roughly, a belief is actionally justified if and only if the believer has applied a procedure that justifies it, while a belief is statally justified just in case the believer *can* apply a procedure that justifies it. Each element of S^B, then, is statally justified so long as I can apply such a procedure to it. And so I can; the procedure consists in an application of any standard set of axioms for number-theory. S^B, therefore, is an infinite sequence of justified beliefs held by the same person.

Does S^B establish the answer "yes" to (QV)? It does so if and only if it is a J-sequence. Hence, given that S^B is an infinite sequence of justified beliefs held by the same person, it establishes the answer if and only if each S^B_i stands in R_1 to S^B_{i+1}. By (DR_1), this amounts to the twofold condition that, first, the propositional object of S^B_{i+1} confirms that of S^B_i and, second, that S^B_i is based on S^B_{i+1}. It is implausible to think that either of these requirements is met. There are now two courses open to someone who wants to show that the answer to (QV) is yes: he can look for a genuine example of a J-sequence or he can try to establish the

[16]Almost the same example is used in Fumerton, pp. 564–65. An objection to examples of this kind has been raised by John N. Williams in "Justified Belief and the Infinite Regress Argument," *American Philosophical Quarterly* 18 (1981), 86; I rebut it in Black, *The Infinite Regress of Justification*, pp. 161–68.

answer other than by appeal to an example. Neither course is promising. I, at least, can think of no potential example more plausible than S^B. As for the second alternative, the procedure would presumably be to argue that a J-sequence *must* exist. But it is hard to envisage such an argument. I conclude that there is no good case for an affirmative answer to (QV).

There is, however, no decisive argument either for the answer no. Someone might for instance seek to establish the answer by reasoning as follows. (1) Suppose the answer to (QV) is yes. (2) Then there is a J-sequence, S^1. (3) By definition of a J-sequence, each S^1_i stands in R_1 to S^1_{i+1}. (4) By the definitions of R_1, of basing and of a causal chain, each S^1_{i+1} stands in the ancestral of . . . causes . . . to S^1_i. (5) Hence each S^1_i is an element of a sequence S^2 such that, first, each element of S^2 is caused by its successor and, second, S^2 is infinite. (6) But there can be no such S^2. (7) So it is false that each S^1_i is an element of such a sequence. (8) So each S^1_i both is and is not an element of such a sequence. (9) The answer to (QV) is therefore no. The controversial step here is the premiss (6). How plausible (6) is depends in part on the force of the "can," that is, on the kind of possibility in question. But there is an intuitive presumption against (6). (6) therefore should be accepted only if there is a powerful argument for it. I know of none and so reject this argument for the answer no to (QV). (A full discussion of (6) would of course involve a review of the different versions of the cosmological argument.)

The best case I can make for the answer is this. I have noted that I cannot think of an example of a J-sequence and that no argument is forthcoming to the effect that a J-sequence *must* exist. In the absence of such an argument, the failure of my search for examples is, assuming I am a competent investigator, itself evidence that no J-sequence exists and hence that (QV) has a negative answer. I tentatively conclude therefore that the answer to (QV) is no, though this conclusion will have to be reversed if anyone finds an example of a J-sequence.

Since (VI) is true just in case the answer to at least one of (QVb), (QVa) and (QV) is no, and since, while (QVb) and (QVa) have affirmative answers, the answer to (QV) is no, (VI) is true. The conclusion that (VI) is true rests thus on the negative answer to (QV) and is reinforced by our intuition in favour of (VI). But the answer to (QV), being tentative, is by itself insufficient to justify a *conviction* that (VI) is true. Nor are the combined weights of the answer and the intuition enough to make such a conviction rational. So the conclusion about (VI), like the answer to (QV), is a tentative one.

Nevertheless it is plausible to think that, if (VI) is after all false, it is falsified only by sequences comprising beliefs of some unusual kind or kinds. The sequence S^B used to establish the answer yes to (QVb) and (QVa) comprises only beliefs about arithmetical propositions. Of course, since (VI) concerns infinite sequences, the class of arithmetical beliefs is the most promising source of examples on which to base affirmative answers to (QVb)–(QV). But it seems likely that, if J-sequences exist, all their elements will fall within this, or within some broader, special domain. In that case the tentative acceptance of (VI) can be turned into a rational conviction that (VI) is true by adding to (VI) a rider preventing its application to those special beliefs that threaten to provide counter-examples to (VI) in its present form.

A simple course is to add a vague clause like "except perhaps in the case of certain kinds of belief.'' But this is unsatisfactory, for it makes (VI) vacuously true; (VI) now says in effect that there are no J-sequences except in cases where there are J-sequences. An informative description must be provided of the category of beliefs to be excluded. A plausible suggestion is that the category consists of beliefs about logic, set-theory and number-theory. I shall not pursue the question whether this description identifies the class of threatening beliefs. Assume that the class has been informatively identified and call its members Z-beliefs. Then a conviction of (VI)'s truth is justified so long as (VI) is modified to state, not that there are no J-sequences, but that if J-sequences exist some of their elements are Z-beliefs.

If (VI) is altered in this way, the last two steps of the infinite regress argument cannot be deduced. To restore the argument's validity it will be necessary also to add to (I) and (II) a rider preventing their application to Z-beliefs and to adjust the other steps accordingly. The result will be that the infinite regress argument concerns the justification only of beliefs that are not Z-beliefs.

This completes the survey of the argument's premisses. To summarize: Premiss (I) is false: it is not the case that a belief is justified only if it stands in R_1 to a justified belief. Premiss (II) is true: justified beliefs exist. Of the two premisses characterizing R_1, (III), which says that R_1 is irreflexive, is true, while (IV), which says that it is transitive, is false. The final premiss, (VI), is true: there are no J-sequences.