Peter Kail’s *Projection and Realism in Hume’s Philosophy* is an excellent book, consisting—like Hume’s *Treatise* itself—of three excellent parts. I will comment on one central aspect of its second part: its explanation of the source of the second thoughts that Hume famously expressed, with a frustrating lack of specificity, about his own initial discussion of personal identity in the *Treatise*.

As is well known, Hume holds in the section “Of personal identity” (T 1.4.6) that a self, mind, or person is “nothing but a bundle or collection of different perceptions” (T 1.4.6.4; SBN 252) and, more specifically, a “system of different perceptions or different existences link’d together by the relation of cause and effect” (T 1.4.6.19; SBN 261). This bundle has neither perfect simplicity (partlessness) at one time nor perfect identity (invarableness and uninterruptedness) through time; nonetheless, he argues, the imagination ascribes both features to it as the result of the associative influence of the relations of causation and resemblance holding among the perceptions themselves. He devotes several pages of the work’s Appendix, published more than a year later in the subsequent volume, to reporting a “difficulty too hard for my understanding” (T Appendix 21; SBN 636) that leads him to despair regarding his previous account, “such a labyrinth that, I confess, I neither know how to correct my former opinions nor how to render them consistent” (T Appendix 10; SBN 633).
There are now well over two dozen proposed explanations in print of the source of Hume’s second thoughts, so the creation of an entirely new and original one is in itself no mean feat. To my knowledge, no commentator on Hume has ever simply endorsed the explanation of another commentator on this question, so it will come as no surprise that I do not endorse Kail’s, just as he does not endorse mine. Nevertheless, I should emphasize at the outset that we are on the same side of what has emerged as an important interpretive divide. This is the divide between, on the one hand, those who interpret Hume as worried primarily about a problem in the *psychology of ascription* by which identity is attributed to the mind; and on the other hand, those who interpret him as worried primarily about a problem in the *metaphysics of bundling* by which perceptions actually belong to or constitute a single mind. Kail and I are both squarely on the latter, metaphysics-of-bundling, side of the divide. This puts us distinctly in the minority, and we are therefore prepared, I believe, to make common cause against a host of shared opponents. Ironically enough, the underlying difference between us is that Kail’s explanation depends on Hume’s taking an attitude toward a doctrine—*Causal Realism*—that my explanation depends on Hume’s *not* taking. This fact may serve as a general warning about what we might fairly call the “interpretive openness” of Hume’s actual text.

In what follows, I won’t say anything more about my own explanation of Hume’s second thoughts, nor will I stop to outline either the section “Of personal identity” or the portion of the Appendix that reconsiders it. Instead, I will begin by listing four criteria, derived from what Hume either says or fails to say in the Appendix, that I believe any successful explanation of the source of his second thoughts should satisfy. 2 Next, I will
summarize Kail’s proposed explanation, which I will call the Realist Explanation. Finally, I will explain why I doubt whether this Realist Explanation satisfies the criteria.

Four criteria that any successful explanation of the source of Hume’s second thoughts should satisfy are as follows:

The Crisis Criterion: The problem ascribed to Hume should be one that he would regard as fitting his initial description of it in the Appendix as one threatening “contradictions and absurdities” and involving him in a “labyrinth” sufficient to inspire skepticism (T Appendix 10; SBN 633).

The Origin Criterion: The problem ascribed to Hume should plausibly fit his own description of when and how the problem arose or was discovered—namely, when proceeding “to explain the principle of connexion, which binds [the mind’s perceptions] together, and makes us attribute to them a real simplicity and identity” (T Appendix 20; SBN 635).

The Solution Criterion: The problem ascribed to Hume should be one that he would regard as fitting his concluding description of it, as involving two principles that he can neither “render consistent” nor “renounce”—namely, the principles “that all our distinct perceptions are distinct existences and that the mind never perceives any real connexion among distinct existences”—and hence also as being at least seemingly soluble if “perceptions inhered in something simple and individual” or “the mind perceived some real connexion among distinct existences” (T Appendix 21; SBN 636).

The Scope Criterion. The problem ascribed to Hume should be one that he would have regarded as applying primarily or especially to his account of personal identity, which is where he locates it—and not, for example, as applying equally and in just the same way to his account of all ascriptions of identity or to his account of mental life more generally.
Kail’s Realist Explanation depends on attributing to Hume the following four doctrines, the formulations of which are derived fairly closely from his book, although the designations given to the first three are of my invention:

**Causal Bundling**: The mind is “a system of different perceptions link’d together by the relation of cause and effect.”

**Separability**: All of our different perceptions are distinct, and all of our distinct perceptions are separable from one another.

**Disconnection**: There cannot be “a real necessary connection” between separable perceptions.

**Causal Realism**: Any causal relation “involves a real necessary connection between cause and effect” (even though such connections may be unknown and even broadly unintelligible to us).

Kail understands a *real necessary connection* between two objects—as referenced in the statement of the last two doctrines—to be a connection the perception of which would show the metaphysical impossibility of one occurring without the other. This is, of course, exactly the kind of perception that Hume has “rejected in all cases” in his earlier discussion of causation (see T 1.4.6.16; SBN 259-260, alluding to T 1.3.14.13; SBN 161-162).

The Realist Explanation of Hume’s problem then proceeds as follows. Hume adopts Causal Bundling at least in part because he rejects the theory that the mind is a mental substance—that is, a substratum in which perceptions inhere. His argument that the mind is not such a mental substance, first offered in the section of the *Treatise* “Of the immateriality of the soul” (T 1.4.5) but partly repeated in “Of personal identity” and again in the Appendix, invokes Separability (T 1.4.5.5; SBN 233; T 1.4.6.3; SBN 252; T Appendix 12; SBN 634). In reviewing that argument in “Of personal identity,” however, he comes to a
realization: although he has been presupposing, as a Causal Realist, the existence of unperceived real necessary connections between the causally-related objects of different perceptions, the separability of those different perceptions shows, by Disconnection, that there cannot be any real necessary connections among the perceptions themselves. Hence, by Causal Realism, there cannot be any causal relations among our different perceptions—a result that violates Causal Bundling.

The ingenuity of this explanation of Hume’s second thoughts is truly admirable. There are several reasons, however, to question whether it satisfies the Crisis Criterion; I will mention two. The first concerns an ambiguity in the formulation of Causal Realism: does it require (a) that no two things can stand in any kind of causal relation unless there is a real necessary causal connection specifically between those two things, or does it require only (b) that no two things can stand in any causal relation unless they are both somehow involved with some real necessary connection or other? This makes a difference, because it is rarely if ever the case for Hume that a perception must be the complete cause, as opposed to a partial cause, of another perception. Typically, other conditions—including, one may reasonably assume, states of the brain—are causally required as well. He holds, for example, that the impression of a cause produces a lively idea of (i.e., a belief in) its effect only in conjunction with a change in background mental organization (specifically, the establishment of a “habit”) that itself has resulted from exposure to a constant conjunction. Similarly, principles of association are only a “gentle force, which commonly prevails” (T 1.1.4.1; SBN 10); but since Hume is a determinist, he is committed to holding that their operation depends partly on other causal factors besides the related perceptions themselves. Now, since a mere part of a complete cause may often occur, in isolation, without the effect
of that complete cause, there can by definition be no real necessary connection between such a partial cause and that effect. Hence, two perceptions can themselves be separable even though they are nevertheless also \textit{causally related} as partial cause and effect. Moreover, this kind of causal relation seems sufficient for Causal Bundling. Thus, reading (a) seems implausibly strong, but reading (b) seems to produce no paradox as long as perceptions are only partial causes of one another.

A second and more basic worry, however, about whether the Realist Explanation satisfies the Crisis Criterion concerns the depth of Hume’s commitment to Causal Realism on either reading of it and regardless of whether perceptions completely or only partially cause other perceptions. There is, of course, a large and lively debate about whether Hume regarded Causal Realism as sufficiently intelligible to be so much as a possible object of a propositional attitude such as belief or presupposition (\textit{presupposition} being Kail’s preferred option). Kail’s book is a major contribution to that debate, and I think he is largely right—with some significant qualifications—about this contentious issue in Humean semantics. Even if Hume regards the doctrine as a \textit{possible} object of belief or presupposition, however, it is doubtful whether he regards it upon final reflection as \textit{probably true}; and it is even more doubtful whether he regards its truth as an essential part of his philosophy, such that its denial would provoke a new skeptical crisis. It is this last level of commitment, however, that the Realist Explanation seems to require—as evidenced by Kail’s subtle slide from the initial formulation “involves a real necessary connection” to “must involve a real necessary connection” once he comes to employ Causal Realism in the explanation of the problem he ascribes to Hume.4

Hume himself, of course, offers a definition of ‘cause’ as follows:

\begin{quote}

\end{quote}
We may define a cause to be “An object precedent and contiguous to another, and where all the objects resembling the former are placed in like relations of precedency and contiguity to those objects, that resemble the latter.” (T 1.3.14.31; SBN; SBN 170)⁵

He certainly does not offer this definition—or, for that matter, the second definition (in terms of association and inference) that immediately follows it—as synonymous with the term ‘cause’. Rather, he is offering, as I understand him, a semantics for causal terms that is derived from what I have elsewhere called his “causal sense” epistemology of causation,⁶ a semantics that parallels the semantics for moral terms he derives from his moral sense epistemology of virtue. Furthermore, this semantics commits him to the view that his first definition, when properly understood, is necessarily co-extensive with ‘cause’, so that its satisfaction—call it “constant conjunction” for short—is necessary and sufficient for the existence of a causal relation.⁷ In particular, should it somehow later emerge that there are or might be cases of constant conjunction without a real necessary connection, his conclusion would be that there are or might be some causes without real necessary connections, and not that there are some constant conjunctions without causation.

This interpretation of Hume’s understanding of the status of the first definition is confirmed by at least three different uses he makes of it. The first use lies in the first two “corollaries” (of four in all) that he draws from his definitions of ‘cause’ immediately after he gives them in the Treatise. These two corollaries are (i) “that all causes are of the same kind” and (ii) that “there is but one kind of [causal] necessity.” In support of the first corollary, he argues:

For as our idea of efficiency is deriv’d from the constant conjunction of two objects, wherever this is observ’d, the cause is efficient; and where it is not, there can never be a cause of any kind. For the same reason we must reject the
The distinction betwixt cause and occasion, when suppos’d to signify any thing essentially different from each other. If constant conjunction be imply’d in what we call occasion, ’tis a real cause. (T 1.3.14.32; SBN 171)

The second corollary, he claims, is established by “the same course of reasoning”—as “appears from the precedent explication of necessity.” This, he writes, is because “’tis the constant conjunction of objects, along with the determination of the mind, which constitutes a physical necessity” (T 1.3.14.33; SBN 171). In the course of discussing this course of reasoning, moreover, he explicitly acknowledges at least the theoretical possibility of “necessary”—and hence presumably complete—causes that nevertheless exhibit some modest degree of “inconstancy” in producing their effects. This would appear to be incompatible with these causes having any real necessary connections, as Kail understands them, with their effects.

Hume uses a similar style of argument in the section of the Treatise “Of liberty and necessity” (T 2.3.1), where he appears to claim that human actions have all the causal necessity that anything can have, and do so solely in virtue of satisfying his two definitions—which, he indicates, cannot really diverge—regardless of any supposed real necessary connection or its absence:

Here then are two particulars, which we are to consider as essential to necessity, viz. the constant union and the inference of the mind; and wherever we discover these we must acknowledge a necessity…. [t]he actions of matter have no necessity, but what is deriv’d from these circumstances. (T 2.3.1.4; SBN 400)

Finally, and of equal importance, he concludes the section “Of the immateriality of the soul” by arguing at length that matter can be the cause of thought precisely because constant conjunction is sufficient for causation. He writes:
All objects, which are found to be constantly conjoin’d, are upon that account *only* to be regarded as causes and effects…. And as the constant conjunction of objects constitutes the very essence of cause and effect, matter and motion may often be regarded as the causes of thought, as far as we have any notion of that relation. (T 1.4.5.32-33; SBN 249-250; emphasis added)

I freely grant that the final phrase of this passage suggests an epistemological qualification that Kail would find significant, but I do not see it as significantly weakening the application of the premise concerning the “very essence of cause and effect.” The theoretical unimportance to Hume of the postulation of real necessary connections is indicated by a remark earlier in the *Treatise*, the evidential weight of which Kail fully grants:

> I am, indeed, ready to allow, that there may be several qualities, both in material and immaterial objects, with which we are utterly unacquainted; and if we please to call these power or efficacy, it will be of little consequence to the world. (T 1.3.14.27; SBN 168)

I conclude that the Realist Explanation does not describe a problem that would have led to a Humean crisis.⁹

Although we have thus far been considering only the Crisis Criterion, the Realist Explanation also runs into difficulty with the Origin Criterion. Kail proposes that Hume realized his problem when, in reviewing the section “Of personal identity,” he came to recognize that his appeal to Separability in arguing against a substantial self was, given Causal Realism and Disconnection, in conflict with Causal Bundling. A statement of Causal Bundling does occur in that section as part of his explanation of the “principles, that unite our successive perceptions in our thought or consciousness.” Even *without* Causal Bundling, however, the conjunction of Separability, Disconnection, and Causal Realism entails, at least according to the Realist Explanation,¹⁰ that there can be no causal relations among different
perceptions. Since Hume’s fundamental philosophical project of a “science of man” demands that there be many such relations, one might have expected this broader problem to have occurred to Hume earlier—especially since the argument against a substantial self originally occurs in the previous section, “Of the immateriality of the soul.” Kail’s response to this worry is that Hume might “perhaps” have been able to recognize this problem earlier, but in fact just did not do so (page 141).11

Equally importantly, however, this difficulty with the Origins Criterion suggests a similar respect in which the Realist Explanation appears not to satisfy the Solution Criterion either. For while granting the inheritance of all of one’s perceptions in a substantial soul (that is, renouncing the first of his two “unrenounceable principles”) would certainly allow Hume to offer an account of the mind’s nature other than Causal Bundling, he would still remain as committed as ever by his project of a science of man to the existence of a multitude of causal relations among the mind’s perceptions. Yet the mere metaphysical dependence of each perception on a common substantial substratum seemingly would not provide any real necessary causal connections among the perceptions themselves of the kind demanded by a Causal Realist construal of that project. In this respect, allowing that “perceptions inhered in something simple and individual” would seemingly not resolve Hume’s more serious underlying problem.

Finally, the Realist Explanation raises the same kind of worry in connection with the Scope Criterion as well. For even if the problem it identifies first occurred to Hume in connection with his statement of Causal Bundling in the section “Of personal identity,” the alleged larger implication of the other three doctrines that there can be no causal relations among perceptions would equally infect the entire science of man, independent of any
theory of the mind’s composition. Kail responds to this worry by claiming that the problem in fact arose in Hume’s own mind in connection with Causal Bundling, and that Hume hoped that a solution to the problem in connection with personal identity would carry over to other instances of it in the science of man as well (page 139). As we have seen, however, a solution in terms of a substantial self—which is, in fact, the kind of solution seemingly most on his mind in the Appendix—would not obviously resolve the larger problem.

I conclude, then, that invoking Causal Realism does not lead to a successful explanation of Hume’s second thoughts. I have only indirectly discussed the case for attributing Causal Realism to Hume, but Kail himself maintains that the ability of such an attribution to explain Hume’s second thoughts in the Appendix “tips the balance” of otherwise equivocal evidence in favor of attributing the doctrine to him. If, however, as Kail concedes, the balance of evidence outside the Appendix for interpreting Hume as a committed Causal Realist is equivocal at best, that makes it all the more unlikely that Hume was nevertheless so deeply committed to the doctrine as to see the problem described by the Realist Explanation as a “labyrinth” leading to skeptical despair. Conversely, if the attempted explanation of Hume’s second thoughts that results from attributing Causal Realism to him is, as I have tried to argue, unsuccessful on other grounds, then Hume’s commitment to Causal Realism—or at least the strength of his commitment—must be brought back into question anew.

It is an ironic feature of the debate over Hume’s recantation of his theory of personal identity that a number of commentators have, over the years, themselves recanted their own explanations of its source. I cordially invite Kail to consider doing the same—while also, of course, reserving the right to recant my own “upon [in Hume’s notable phrase] more mature reflection.”
NOTES

I am grateful to Angela Coventry for organizing the Hume Society group session at the 2010 American Philosophical Association Central Division Meetings at which this paper was originally presented. I am grateful to Peter Kail, Eric Schliesser, Jacqueline Taylor, Angela Coventry and the audience of that session for helpful comments and discussion. I am also grateful to Galen Strawson for comments on a written version of the paper.


2 As indicated in “Rethinking Hume’s Second Thoughts About Personal Identity,” op. cit., I would also accept a Difficulty Criterion: It would be at least desirable if the problem were one that, while within Hume’s ability to notice upon review, would also be at least somewhat difficult for him to express clearly. For despite his enormous gifts as a writer, he does not manage to say what it is. I will not invoke this fifth criterion against Kail’s explanation, however.

3 More specifically, I read Hume as holding that we typically “presuppose” the existence of real necessary causal connections not by formulating any idea of such connections, but rather by conflating two distinct species of necessity: the demonstrative (and intrinsic) necessity pertaining to relations of ideas and the causal (but non-intrinsic and associative) necessity that requires constant conjunction. This conflation leads us to reason as though we perceived one necessity that is both demonstrative and causal.

I think that Hume firmly rejects the legitimacy of this kind of “presupposition” because the attempt to formulate a corresponding belief involves the presupposer in a forced choice between contradiction and meaninglessness. I also think, however, that Hume allows that the mechanism of “relative ideas” permits—barely—the formulation of what Kail calls the “Bare Thought” of something necessity-related that applies to causes and effects but is intrinsic to them rather than a product of association. This idea must be doubly-relational, however: it is either (i) the idea of some unknown quality in objects that participates in some relation sufficiently similar to known demonstrative and causal necessities to count as a third kind of necessity; or (ii) the idea of some unknown quality of objects that can be represented by a kind of thing sufficiently similar to known ideas to count as a second kind of idea that participates in unknown but demonstrative “relations of ideas.” The formation of such a “relative idea” is rare in either form, Hume thinks, but understandable through natural principles of the mind as a response to recognizing the illegitimate conflation. Because it is an idea (albeit a relative one), it can be the object of belief as well as of presupposition without doing violence to Hume’s theory of belief as a “lively idea.” For further explication and defense, see my article “Hume” in The Oxford Handbook of Causation, edited by Helen Beebee (Oxford: Oxford University Press, 2009): 73-91.

4 Compare, for example, Projection and Realism in Hume’s Philosophy pages 125 and 132 with pages 140 and 142.

5 He subsequently limits the requirement of spatial contiguity to those objects that stand in spatial relations—as desires and smells, for example, do not.

6 See my “Hume,” op. cit.

7 Indeed, I would argue that this is also true of the second definition, on his view. See “The Representation of Causation and Hume’s Two Definitions of Cause,” Noûs 27.2 (June 1993): 167-190; Cognition and Commitment in Hume’s Philosophy, op. cit. Chapter 5; and “Hume,” op. cit.

8 For Kail’s argument that Hume’s treatment of “liberty and necessity” is compatible with Causal Realism, see his “How to Understand Hume’s Realism,” in The New Hume Debate, 2nd edition, edited by Rupert Read and
perceptions might fail to be adequate representations of them—especially when it seems that these ideas differ having them. Hence it may seem un-Humean even to consider the prospect that our ideas of our own (137n); but I’m afraid I don’t see quite how these appeals work. To be sure, Hume also affirms—in passages “phenomenal distinctness” and by an appeal to an alleged “vicious regress” about secondary perceptions (i.e., perceptions of other perceptions), and he replies (if I understand him properly) both by appeal to similarities of specific degree of force and vivacity) is ever actually found to occur without one like the first. Hume’s general test for whether any two unseparated things are nevertheless separable is whether they can be separated, as he says, “in the thought or imagination” (T 1.1.7.3; SBN 18)—in other words, whether ideas of them do occur separately (either naturally or as a result of specific volition). This “conceivability” test would presumably be his test for the separability of never-actually-separated perceptions as well. Underwriting this test, of course, is his Conceivability Principle: “Whatever we conceive is possible, at least in a metaphysical sense” (Abstract of ... A Treatise of Human Nature 11; SBN 650).

Now, Causal Realism is prima facie incompatible with the Conceivability Principle as just stated, inasmuch as Causal Realism demands that the various causes and effects, of all kinds, that Hume readily grants we do conceive to be separated must nevertheless not be metaphysically separable. As Kail quite properly emphasizes, however, Hume often limits the scope of the Conceivability Principle to things that we can conceive “clearly and distinctly”—and this restriction allows Causal Realists to reply that we do not have ideas of causes and effects that are sufficiently clear and distinct as representations of their objects to license our applying the Conceivability Principle to them. Yet once this kind of doubt arises about our ability to apply the Conceivability Principle to causes and effects generally (as the Causal Realist demands that it does), it seems we may also readily begin to doubt whether any of our causally related perceptions really are metaphysically separable in fact, even though they seem separable in conception.

To be fair, Kail considers a form of this objection, even noting that Hume recognizes “secondary perceptions” (i.e., perceptions of other perceptions), and he replies (if I understand him properly) both by appeal to “phenomenal distinctness” and by an appeal to an alleged “vicious regress” about secondary perceptions (137n); but I’m afraid I don’t see quite how these appeals work. To be sure, Hume also affirms—in passages that Kail also cites—that we cannot be mistaken about the intrinsic character of our perceptions while we are having them. Hence it may seem un-Humean even to consider the prospect that our ideas of our own perceptions might fail to be adequate representations of them—especially when it seems that these ideas differ from the primary perceptions that are their objects only or chiefly in some respect like Humean “force and vivacity.” Nevertheless, if Hume really is committed to the view that all causally-related objects are related by unknowable-in-principle real necessary connections, as Causal Realism demands, then it is at least not obvious why he should be required to regard the existence or non-existence of such unknowable relational characteristics of perceptions as falling within the scope of the transparency of their intrinsic features—for there can be no doubt that he allows errors about merely relational features of ideas generally.

9 A further problem for the satisfaction of the Crisis Criterion arises even if perceptions are regularly the complete causes of other perceptions. It concerns the question of whether a commitment to Causal Realism would not simply lead Hume to reject or revise Disconnection (i.e., the impossibility of real necessary connections between perceptions) as a matter of course, without experiencing any crisis or despair. After all, any two particular perceptions that occur together are, ex hypothesi, not actually separated; and if the first is ever to be properly regarded as the complete cause of the second, then it also follows for Hume, by his first definition of ‘cause’, that no other perception relevantly like the second (which may of course include similarities of specific degree of force and vivacity) is ever actually found to occur without one like the first. Hume’s general test for whether any two unseparated things are nevertheless separable is whether they can be separated, as he says, “in the thought or imagination” (T 1.1.7.3; SBN 18)—in other words, whether ideas of them do occur separately (either naturally or as a result of specific volition). This “conceivability” test would presumably be his test for the separability of never-actually-separated perceptions as well. Underwriting this test, of course, is his Conceivability Principle: “Whatever we conceive is possible, at least in a metaphysical sense” (Abstract of ... A Treatise of Human Nature 11; SBN 650).

10 I enter this qualification because of the issue described earlier concerning the interpretation of Causal Realism and the potential partiality of the perceptual causes of perceptions.

11 Kail also suggests that, in “Of the immateriality of the soul,” Hume was primarily thinking of perceptions in terms of their representational content rather than their own existence as mental entities. This seems unlikely to me, given that the section is concerned specifically with the possible “local conjunction” of perceptions with one another and the suggestion that they are themselves ontologically dependent on a mental substratum.