

# 8

## Personal Identity and the Skeptical System of Philosophy

Corliss Gayda Swain

The discussion of personal identity occupies so few pages of the *Treatise* and its Appendix that mastering Hume's views on this topic may appear easy. This appearance would, however, be deceiving. To understand Hume's account of personal identity one must, as Hume says, "take the matter pretty deep" (T 1.4.6.5; SBN 253). Because Hume, often implicitly, but sometimes explicitly, compares his views on personal identity with his views on other metaphysical issues, especially causation and external objects, one cannot achieve a full understanding of Hume's theory of personal identity without understanding those views, the discussion of which takes up much of Book I of the *Treatise*. Understanding Hume on these other matters is also necessary for understanding the puzzling comments Hume makes in the Appendix about his theory of personal identity. While this approach requires considerable effort, the effort is well worth it: for it leads to a deeper and more unified understanding of Hume's skepticism – its nature, grounds, and limits.

Interpretations of Hume's skepticism have evolved. Traditionally, Hume was thought to be skeptical of causation, external objects, and persons, and it was thought that Hume wanted to demonstrate the irrationality of our everyday beliefs about these things. The interpretation I offer below continues the reassessment of that tradition, a reassessment traceable to Norman Kemp Smith's *The Philosophy of David Hume* (1941). New interpretations take Hume's skepticism to be far more limited than the tradition suggests, aimed more at philosophical views of causation, external objects, and the self than at these realities themselves. (See Strawson 1989; Baier 1991; Garrett 1997.)

Influential representatives of the traditional interpretation of Hume's theory of personal identity include Thomas Reid, Terence Penelhum, Roderick Chisholm, and Derek Parfit. Robert Fogelin and Wade Robison hold intermediate views. Fogelin (1985) attempts to reconcile Hume's naturalism with his skepticism, but he nonetheless sees Hume as an extreme skeptic. Robison (1974) interprets Hume as explaining the content and origin of popular beliefs about the self while denying the truth of these beliefs. The traditional interpretation is often presented as uncontroversial in discussions of personal identity that include an historical survey. See for example, Noonan 2003: ch. 4 and Olson 1997: ch. 7. However, most Hume scholars today see Hume as offering a positive account of the enduring self (e.g. Passmore, Pike, Noxon, Biro, Stroud, MacIntyre, Beauchamp, Bricke, Traiger, Flage, Swain, Loeb, Waxman, and Baxter to name only a few). Their disagreements concern the nature and adequacy of Hume's account.

Hume's views on personal identity are not difficult to summarize. The first element of his view is negative: Hume rejects a common philosophical view of the self according to which the self has both perfect simplicity and perfect identity. (Hume uses the terms *self*, *person*, and *mind* interchangeably, and in this article I follow his practice.) A being has *perfect simplicity* just in case it is not composed of parts at any given time; it has *perfect identity* just in case it is composed of exactly the same parts from one time to the next. When articulating his view of personal identity, Hume distinguishes between objects that have perfect identity and other objects that continue in existence over time. To illustrate the notion of perfect identity he uses the example, familiar from Locke (1975: 2.27.3), of a mass of particles. Such a mass becomes a different entity when and only when particles are subtracted from the mass or added to it. Whether they are simple or complex, whether made up of bits of matter or composed of something immaterial, only those beings whose composition remains unchanged through time have what Hume calls "perfect identity" (T 1.4.6.8; SBN 255–6). Hume denies that the self has this kind identity. Instead, he argues that the identity we attribute to the self is an "imperfect identity," the same kind of identity that we attribute to other objects whose parts change over time, including plants, animals (T 1.4.6.5; SBN 253 and T 1.4.6.12; SBN 257), ships (T 1.4.6.11; SBN 257), and republics (T 1.4.6.19; SBN 261). Even though we think of these things as unified, singular entities, continuing in existence over time, none of them has "perfect simplicity" or "perfect identity." Hume's positive view, then, is that the self is a composite (T 1.4.6.22; SBN 263), and, regarding its identity, Hume holds that the parts that constitute the self do not remain the same throughout a person's existence, but "succeed each other with an inconceivable rapidity . . . in a perpetual flux and movement" (T 1.4.6.4; SBN 252).

On Hume's positive theory, the parts that make up the self, or mind, are perceptions. These perceptions include not only what Hume calls "IMPRESSIONS OF SENSATION" (sensory perceptions) and "IMPRESSIONS OF REFLECTION" (passions and emotions), but all other kinds of conscious mental states as well (memories,

ideas, thoughts, and occurrent beliefs). (See 10: THE POWERS AND MECHANISMS OF THE PASSIONS.) The mind or self is not, according to Hume, something distinct from its thoughts, emotions, and sensory experiences. It is, instead, a composite comprising all these things.

Commentators disagree about the proper interpretation of the passages in which Hume criticizes the view that the self is simple and unchanging (Traiger 1985 and Ainslie 2004 summarize alternative interpretations). We can, nonetheless, find in these passages at least three arguments for the claim that the self is composed of its perceptions. One argument is that we could make no sense of the fact that one's perceptions *belong* to one's self given the further fact, insisted on by Hume, that each perception is a distinct being that is capable of existing on its own, unless we granted that these distinct perceptions *constitute* the self as parts of the whole (T 1.4.6.3; SBN 252). A second argument rests on two premises: (1) that we can never perceive our own minds except by perceiving our thoughts, sensations, and emotions; and (2) that thoughts, sensations, and emotions are *all* we perceive when we perceive our own minds. Proof that we cannot perceive our minds except by perceiving thoughts, sensations, and emotions is supposedly found in the fact that “[w]hen my perceptions are remov’d for any time, as by sound sleep; so long am I insensible of *myself*” (T 1.4.6.3; SBN 252). Finally, the constitutive relation between one's perceptions and one's self is supposed to be evidenced by the fact that “were all my perceptions remov’d by death, and cou’d I neither think, nor feel, nor see, nor love, nor hate after the dissolution of my body, I shou’d be entirely annihilated” (T 1.4.6.3; SBN 252). These arguments provide some grounds, though perhaps not conclusive reasons, for believing that there is no such thing as a mind that is completely devoid of mental contents and that there is nothing to the mind over and above its contents.

Hume sometimes describes the self as a “bundle or collection of different perceptions” (T 1.4.6.4; SBN 252), and even as a “heap” (T 1.4.2.39; SBN 207). For this reason some commentators have called Hume's theory a “bundle theory of the self” (see Pike 1967; Patten 1976; and Baier 1979). However, to the extent that bundles and heaps are arbitrary aggregates of parts, this name is somewhat misleading. Hume also describes the self as a system. He claims that “the true idea of the human mind” is an idea of “a system of different perceptions or different existences, which are link’d together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other” (T 1.4.6.19; SBN 261). Instead of calling his positive theory a bundle theory, we might less misleadingly characterize it as a causal theory of the self. Hume holds that the self's continuance through time is a matter of its being organized as a system. This is not to deny that the self continues in existence through time. It simply affirms that the continued existence of the self, like the continued existence of almost any organized system, including that of organisms, does not depend on the continued existence of the parts that constitute that system at any moment. While Hume calls the identity of all such changing objects “fictitious” (T 1.4.6.15; SBN 259), he

opposes fictitious identity not to real identity but only to perfect identity. (Swoyer 1982 and Traiger 1987 discuss Hume's use of the term "fiction" in his accounts of identity.)

Although Hume does not mention him specifically in this discussion, his theory of personal identity owes much to Locke. Locke had already made the distinction between the kind of identity we attribute to unchanging objects like the mass of matter and the kind of identity we attribute to objects whose constituents change over time (1975: 2.27.4–6). Locke also developed a systemic or causal theory of identity to account for the identity of plants and animals (2.27.4–6). Hume accepts Locke's account of the identity of plants and animals, and he believes there is "a great analogy betwixt [that identity] and the identity of a self or person" (T 1.4.6.5; SBN 253). Hume does not, however, accept Locke's account of personal identity. Locke believed that memory links together the distinct experiences that make up the self, forming a single consciousness that stretches over time (1975: 2.27.9–10 and 2.27.17–23). Hume rejected this hypothesis about the connecting principle, arguing as follows:

Who can tell me, for instance, what were his thoughts and actions on the first of *January* 1715, the 11th of *March* 1719, and the 3d of *August* 1733? Or will he affirm, because he has entirely forgot the incidents of these days, that the present self is not the same person with the self of that time; and by that means overturn all the most establish'd notions of personal identity? (T 1.4.7.20; SBN 262)

Relying on "established notions" of personal identity, according to which the self – not just the man, as Locke had claimed (1975: 2.27.20) – extends beyond its own memories, Hume proposes that the perceptions that make up the self are linked not by memory alone but by many different causal relations.

On Hume's theory, memory is one connecting principle, insofar as it causes new perceptions that resemble earlier perceptions (T 1.4.6.18; SBN 260–1), and resemblance as well as cause and effect can produce the mental associations that are sources of beliefs about the enduring self. However, memory is not *the* connecting principle linking the perceptions that make up the self. Once causal laws governing our psychology have been discovered, we can "extend the same chain of causes, and consequently the identity of our persons, beyond our memory, and can comprehend times, and circumstances, and actions, which we have entirely forgot, but suppose in general to have existed" (T 1.4.6.20; SBN 262). By taking into account causal connections between mental states in addition to those involved in memory, Hume's theory can accommodate the fact that some irremediably forgotten experiences are considered as much a part of one's self as the experiences one remembers.

Hume first restricts his attention to "personal identity, as it regards our thought or imagination" (T 1.4.6.5; SBN 253) when looking for the relation that grounds our judgments of identity. There he finds a number of causal relations between

perceptions: “Our impressions give rise to their correspondent ideas; and these ideas in their turn produce other impressions. One thought chases another, and draws after it a third, by which it is expell’d in its turn” (T 1.4.6.19, SBN 261). When Hume expands his view beyond the thought and imagination, considering, in addition, “our passions and the concern we take for ourselves” (T 1.4.6.5; SBN 253), he finds further evidence in support of his causal theory of the self. Echoing Locke once again (1975: 2.27.23, 25–6), Hume says, “our identity with regard to the passions serves to corroborate that with regard to the imagination, by the making our distant perceptions influence each other, and by giving us a present concern for our past or future pains or pleasures” (T 1.4.6.19; SBN 261). These newly considered causal connections between mental states provide further reasons for rejecting a memory account of personal identity: For it is obvious that the causal connections between present decisions and future happiness, in the case of prudential thoughts, and between past actions and present misery, in the case of regrets, are not produced by memory alone (T 1.4.6.20; SBN 261–2). (For other views concerning the relation between Hume’s view of personal identity with respect to the understanding and with respect to the passions see Capaldi 1989; Henderson 1990; Baier 1991; ch. 6; McIntyre 1989 and 1993; Chazan 1992; and Purviance 1997.

According to Hume, “the true idea of the human mind, is to consider it as a system of different perceptions . . . which are link’d together by the relation of cause and effect” (T 1.4.6.19; SBN 261). He thus seems to embrace a psychological connectedness theory of personal identity. However, Hume’s theory of personal identity is complicated by the further fact that, according to him, “the union of cause and effect . . . resolves itself into a customary association of ideas” (T 1.4.6.16; SBN 260). This has important implications.

Hume asks two related questions about the identity that we attribute to persons: “whether [this relation of identity] be something that really binds our several perceptions together or only associates their ideas in the imagination” (T 1.4.6.16; SBN 259) and “whether in pronouncing concerning the identity of a person, we observe some real bond among his perceptions, or only feel one among the ideas we form of them” (T 1.4.6.16; SBN 259). His answer is informed by his theory of causation. Since he accepts the principle that “the understanding never observes any real connexion among objects” (T 1.4.6.16; SBN 259–60; and T Appendix 20; SBN 635), including connections between causally related objects, Hume concludes that “identity is nothing really belonging to these different perceptions, and uniting them together; but is merely a quality, which we attribute to them, because of the union of their ideas in the imagination, when we reflect on them” (T 1.4.6.16; SBN 260; and T Appendix 20; SBN 635). For Hume, the link between the perceptions that constitute the self is to be found in the observer, who connects together her ideas of those perceptions, rather than where we would expect it – in the observed person, where it would connect the perceptions that actually constitute the self.

Although he concedes that with his causal theory of the self, there are no known – or knowable – connections among the perceptions that constitute a self, Hume does not take judgments of personal identity to be arbitrary and unjustified, since those judgments must be grounded in real, observed regularities. On Hume’s theory of the self, if the mind’s activities lacked the regularities that ground causal judgments when observed, we could make no correct attributions of personal identity. Because the mind’s activities do exhibit the kind of regularity that grounds causal judgments, we can justifiably and non-arbitrarily connect ideas of the perceptions that make up the self when we reflect on them.

Correctly characterizing Hume’s theory of personal identity is not easy. Although we earlier called his theory of personal identity a causal theory, we may, on reflection, decide that Hume does not actually offer an account of personal identity because he never explains what connects the connected perceptions that constitute a mind. If the self is a composite, and its parts “form a whole only by being connected together” (T Appendix 20; SBN 635), then it seems reasonable to insist that a satisfactory account of the self explain what connects these parts. Because Hume argues that no such explanation is possible, his account could be characterized as a “no-theory” or skeptical account of personal identity.

Hume’s views on personal identity are like his views on causation in this respect. Just as we may be tempted to say that Hume has no metaphysical account of causal connections because he believes that we have no idea of the connecting principle that binds together every cause and effect, so also we may be tempted to say that Hume has no metaphysical account of the self because he believes that we have no idea of the connecting principle that binds together the perceptions that make up a self. In one sense that is true, but in another sense it is quite far from the truth.

Consider the case of causation: Some philosophers would say that Hume does have a coherent, novel, and even plausible metaphysical account of causation. However, for Hume, the causal link is not a link between the causally related objects; it is, instead, a link in perceivers’ minds between their perceptions of the causally related objects. Cause and effect are not directly connected; instead, perceptions – ideas and impressions – of causes and effects are connected in the minds of perceivers. If we insist that a satisfactory metaphysical account of causal relations must identify a direct bond between the causally related objects, or that any proposed link must be a link between the causally related objects themselves, we must deny that Hume offers a satisfactory metaphysical account of causation. From this perspective, Hume appears to be a skeptic about metaphysical accounts of causation (though not necessarily about causation itself). The grounds of this skepticism are that *cause and effect are distinct existences*, and that *the mind never perceives connections among distinct existences*. We could, however, with equal justification insist that a satisfactory metaphysical account of causation must explain what the actual connection is between causally connected objects. If the connection is, in fact, an indirect connection that exists only in the minds of perceivers when they

observe (or think about) the causally related objects, then an account that says this would satisfy the actuality requirement despite its failure to say what directly connects the supposedly connected objects. Viewing matters from both perspectives enables us to see how Hume could be both a skeptic and a naturalist when it comes to causation. He is skeptical of metaphysical accounts of causal connections, but such skepticism is compatible with belief in causation and (indirect) causal connections.

Returning to Hume's account of personal identity, if we insist that a satisfactory metaphysical account of the self must explain what actually binds together the perceptions that constitute a self, then we ought to take Hume to be arguing for skepticism about metaphysical accounts of the self. His grounds are very like the grounds appealed to in connection with causation: "*that all our distinct perceptions are distinct existences and that the mind never perceives any connections among distinct existences*" (T Appendix 21; SBN 636). However, if we allow that a satisfactory metaphysical account of the self need only identify the actual connection between the perceptions that make up the self, and if that connection is one that exists only in the minds of perceivers when they reflect on the constituting perceptions, then an account that says this would be a satisfactory metaphysical account according to the actuality requirement.

The words we use to characterize Hume's account – whether we call it a *satisfactory metaphysical* account or not – are unimportant, but there is something at stake. Whether we accept a proposed account of causation or the self depends on what constraints we impose on such accounts. Hume found himself forced to reject some plausible constraints because he believed they could never be satisfied. In explaining why they cannot be satisfied, Hume provided grounds for a limited and principled skepticism. His grounds rule out the possibility of accounts that try to explain connected phenomena in terms of connecting principles. (We could call these "metaphysical glue" accounts.) After rejecting these constraints, Hume proposes alternative accounts of causation and the self which do not even attempt to supply the metaphysical connecting principle. But if the only defect of these accounts is that they do not satisfy the a priori constraints, and if Hume has proved that these constraints cannot be satisfied anyway, then one might reasonably conclude that the constraints themselves, rather than Hume's alternative accounts, should be rejected. (A version of this argument is presented in Swain 1991.) This is a powerful sceptical conclusion concerning the possibility of metaphysics and its scope.

Hume's account of personal identity involves three main ideas: (1) that the self is a composite with the kind of identity – an imperfect identity – shared by all objects that are not composed of exactly the same constitutive parts throughout their existence, (2) that the self forms a whole over time by being composed of successive parts linked together by causal relations, and (3) that the only connection we can observe (or even conceive of) with respect to the parts that successively make up the self is a mental connection between, not the parts themselves, but

only our ideas of these parts, a connection forged when reflecting on perceptions that exhibit causal regularities. Hume's account, or at least this set of three ideas, is consistent. This account nevertheless has an air of paradox, since it affirms that the self is a system of connected perceptions but goes on to say that the only knowable link between the supposedly connected parts is a link between *ideas* of these parts rather than a link between the parts themselves.

Hume's account of the self shares this feature with his account of causation. Hume characterizes his views concerning a necessary connection between cause and effect as paradoxical (T 1.3.14.24; SBN 166), because he accepts the paradoxical conclusion that the "connexion and necessity" of causes "belongs entirely to the soul, which considers the union of two or more objects in all past instances" (T 1.3.14.23; SBN 166). This "paradox" perfectly mirrors the "paradox" concerning the self outlined above. Hume agrees that causes and their effects are necessarily connected, but he observes that the only link he can discover between these connected objects is a link between our *ideas* of the objects rather than a link between the objects themselves. While Hume's account explains why we justifiably think of certain objects as necessarily connected, it does so by explaining our grounds for connecting the *ideas* of these causally related objects (namely, their regular conjunction) rather than by identifying the link that connects causes to their effects and explaining our epistemic access to this link. Hume anticipates his critics' complaining that he would "remove [the causal power] from all causes, and bestow it on a being that is no ways related to the cause or effect, but by perceiving them" (T 1.3.14.26; SBN 168), and sees them ridiculing his position as "a gross absurdity, and contrary to the most certain principles of human reason" (T 1.3.14.26; SBN 168). Hume also foresees his critics' accusing him of rejecting the truism that "[t]hought may well depend on causes for its operation, but not causes on thought" (T 1.3.14.26; SBN 167). Although he is well aware of these paradoxical features of his account of causation, Hume nonetheless offers this account as a true account. Given that it *is* paradoxical, the only hope Hume says he has of its acceptance is "by dint of solid proof and reasoning" (T 1.3.14.24; SBN 166).

When assessing Hume's accounts of the self and of causation, it is important to note that, for Hume, the materials out of which a unified system is constructed exist and are related to each other in definite ways prior to the mind's synthesizing activities. On Hume's view, the mind does not invent fire, wood, or ashes. Nor does it invent the constant conjunction expressed in the proposition, "Ashes are the product of burning wood." This much is perfectly objective. But insofar as the causal connection is supposed to involve more than constant conjunction, there is an extra bit – the necessary connection – which is, according to Hume, a product of the mind's synthesizing activity. Philosophers who take Hume's account of causation to be a simple regularity theory deny that there is more to Humean causation than constant conjunction. However, while that interpretation is consistent with Hume's first definition of a cause, it does not comport well with

Hume's second definition, which refers to the synthesizing activity of the mind: the way that ideas are united in the imagination (T 1.3.15.35; SBN 172). To do justice to both definitions, we should take Hume's theory of causation to involve both an objective element that consists of facts about regularities and a subjective element, consisting of facts about the mind's synthesizing activities. (See 6: CAUSATION for discussion of Hume's two definitions.)

Similarly, on Hume's view of the self, the thoughts, sensations, and emotions which constitute the self exist independently of any synthesizing activity of the mind. These thoughts chase, draw after, and expel each other (T 1.4.6.19; SBN 261) with the regularity characteristic of causally connected things prior to and independently of reflection. If they did not, we would have no grounds for connecting our ideas of these perceptions together. Persons are thus not *completely* self-made. But they are self-made to some extent: For although reflection discovers, rather than produces, the thoughts and their regular conjunctions, reflection on mental contents and their regularities does produce a sense of a unified and continuing self by producing the feeling that these conjunctions are necessary and that the conjoined items are connected.

Interpreters who take Hume's theory of causation to be a straightforward regularity theory tend to ignore what Hume says about the synthesizing activities of the mind and the subjective side of causal connections. With interpretations of Hume's view of the self, we find a tendency to focus exclusively on the synthesizing activities of the mind or the subjective side of self-construction, while the objective regularities that ground this activity tend to get ignored. The resulting view is that persons are almost completely self-made, and the only facts about what belongs to any given mind are those that depend on the synthesizing activity of the mind. To do justice to Hume's account of the self, however, we must recognize that the possibilities of self-construction are heavily constrained by the stubborn facts of human psychology.

Any assessment of Hume's theory of the self must take into account Hume's remarks in the Appendix to the *Treatise*. In the Appendix, Hume was concerned to clarify his positions and to correct any mistakes he had made. He identifies only three mistakes, one of which he characterizes as a "considerable mistake" (T Appendix 1; SBN 623), while the other two are called "errors of less importance" (T Appendix 22; SBN 636). The considerable mistake appears to implicate Hume's theory of personal identity. The section of the Appendix where Hume discusses this mistake begins as follows:

I had entertain'd some hopes, that however deficient our theory of the intellectual world might be, it wou'd be free from those contradictions, and absurdities, which seem to attend every explication, that human reason can give of the material world. But upon a more strict review of the section concerning *personal identity*, I find myself involv'd in such a labyrinth, that, I must confess, I neither know how to correct my former opinions, nor how to render them consistent. (T Appendix 10; SBN 631)

This passage admits of two interpretations. The familiar interpretation holds that Hume discovers a mistake in his theory of personal identity. On a different interpretation, which I develop below, Hume's considerable mistake does not involve the theory of personal identity; instead, the mistake involves Hume's earlier claim that accounts of the intellectual world (or the self) are special in that they alone are, or can be, free of contradictions. (Fieser [1989:105–6] and Waxman [1992: 233] also suggest that the passage in the Appendix refers to Hume's claim about the intellectual world in T 1.4.5.1.)

In the Appendix passage, Hume appears to claim that his theory of personal identity involves unavoidable inconsistencies. Because finding unavoidable inconsistencies in one's theory is generally held to be a bad thing, most commentators who have grappled with Hume's theory of personal identity have interpreted this passage as evidence that Hume was dissatisfied with his earlier account and had, upon review of it, found reasons for rejecting it. Once they accept that Hume found something wrong with his account of personal identity and *only* in his account of personal identity, since he mentions only this section, their next task is to identify the precise source of Hume's dissatisfaction. (Both Stroud [1977: ch. 6] and Pears [1990: 120] argue persuasively that the problem must be unique to Hume's account of personal identity.) This has proved difficult.

Commentators agree that if Hume did find some new problem when he reviewed his section on personal identity, he wasn't forthcoming about its nature in the Appendix. Hume mentions inconsistencies, but the supposedly inconsistent principles that he refers to are logically compatible and are the grounds for his own skeptical accounts of the self, external objects, and causation. Hume also points out as a deficiency in his account of personal identity the fact that it does not provide the principle that connects the perceptions that make up the self, but this deficiency was already remarked in the *Treatise*. It is not something that would have struck Hume only after he reviewed his section on personal identity. While much ingenuity and philosophical acumen has been exhibited by commentators who believe Hume did find some new problem with his account of personal identity, none of the many attempts to locate the problem and to show that it infects only Hume's account of personal identity has won widespread acceptance. See Stroud 1977: ch. 6; Garrett 1981 and 1997: ch. 8; and Fogelin 1985: ch. 8 for overviews and criticisms of various interpretations. Stroud, Garrett, and Fogelin also offer their own interpretations, but none has gone unchallenged.

Given the difficulty of finding a problem that infects only Hume's account of personal identity and fits with what Hume says in the Appendix, we have reason to welcome an alternative interpretation of Hume's "considerable mistake." On this interpretation, the mistake is not found in Hume's section on personal identity but rather in the statement he made in the immediately preceding section, "Of the immateriality of the soul." That section begins by drawing a contrast between accounts of the external world – accounts that explicate our idea of matter – and accounts of the mind, which explicate our idea of the self. Hume says the former

contain “contradictions” as well as “difficulties” (T 1.4.5.1; SBN 232), while the latter, although they involve difficulties, need contain no contradictions. He draws the contrast as follows:

Having found such contradictions and difficulties in every system concerning external objects, and in the idea of matter . . . we shall naturally expect still greater difficulties and contradictions in every hypothesis concerning our internal perceptions, and the nature of the mind . . . But in this we shou’d deceive ourselves. *The intellectual world, tho’ involved in infinite obscurities, is not perplex’d with any such contradictions, as those we have discover’d in the natural. What is known concerning it, agrees with itself;* and what is unknown, we must be contented to leave so. (T 1.4.5.1; SBN 232; my italics)

This is the mistaken claim that Hume is concerned to correct when he says, “I had entertain’d some hopes, that however deficient our theory of the intellectual world might be, it wou’d be free from those contradictions, and absurdities, which seem to attend every explication that human reason can give of the material world” (T Appendix 10; SBN 633). Interpretations of the significance of Hume’s mistake will differ depending on how Hume’s claims about contradictions are understood. Fieser (1989) and Waxman (1992), for example, take the contradictions to signal something quite different from what I propose. We nonetheless agree that “strict review of the section concerning *personal identity*” (T Appendix 10; SBN 631) made Hume realize that his claim about the possibility of a consistent, though incomplete, metaphysical account of the “intellectual world” was mistaken. The interpretation I present below argues that the contradictions Hume refers to teach us only about the futility of trying to give certain kinds of metaphysical accounts; they do not indicate that there is some special problem with Hume’s positive account of personal identity. (For an alternative interpretation, see Winkler 2000.)

On my interpretation, the vexed and seemingly unanswerable question, “What new problem concerning his theory of the self was Hume alluding to in the Appendix?” need not be answered, since on this interpretation there is no reason to think that Hume *did* allude to some new problem in his account. Rather than looking for new problems, we can focus, instead, on the more general and important task of understanding the kinds of contradictions and absurdities that Hume believes infect every attempt to account for our most central and important beliefs: What are the contradictions? Why do these contradictions lead to skepticism? and What are we supposed to be skeptical about?

The only “inconsistency” or “contradiction” that Hume mentions in the Appendix is between two principles: “*that all our perceptions are distinct existences, and that the mind never perceives any real connexion among distinct existences*” (T Appendix 21; SBN 636). These principles play an important role in Hume’s accounts of causation and of external objects as well as in his account of the self.

It is obvious that the principles are not inconsistent with each other. But if they are both true, the implications are far-reaching and radical. If they are true, then whenever we find distinct existences, even if we have good reasons to believe that they are connected so as to form some sort of system, we will be frustrated in our attempts to discover the links that actually connect them. And we find distinct existences in very important places. Hume points out that causes and effects are distinct existences that we nonetheless take to be connected. The regularity of their conjunction is our grounds for believing that there is a connection, but we can find nothing in the causally related objects – or anywhere else – that would explain this regularity. Hume argues that anyone who has tried to give an account of causal powers or to identify the connecting link between cause and effect has run afoul of one or the other of the two principles: either by failing to acknowledge the distinctness of cause and effect, or by speaking of some sort of metaphysical bond without having any idea of the bonding agent (T 1.4.3.9–10; SBN 223–4). In Hume’s words, “it appears, that when we say we desire to know the ultimate and operating principle, as something, which resides in the external object, we either contradict ourselves, or talk without a meaning” (T 1.4.7.5; SBN 267).

Hume’s discussion of our idea of external bodies has a similar structure. He describes our belief that objects persist over time as a belief about connected objects or perceptions. To believe that the desk I see now is the same as the desk I saw earlier is simply to believe that there is a special kind of connection between the desk I see now and the desk I saw earlier. Hume does not take this belief to be arbitrary or unjustified. The observed coherence and consistency of our experiences of external objects constitute our grounds for believing that the objects are connected. However, we can find nothing – in the connected object themselves or anywhere else – that would explain this coherence and consistency or provide the principle of connection. Those who have tried to give an account of the connecting principle (linking together the parts that constitute the enduring objects of perception) have run afoul of one or the other of the two principles: either they fail to recognize the distinctness of the distinct objects of perception, which is the case with the vulgar (T 1.4.2.43; SBN 209), or they speak of some sort of metaphysical bonding agent (like a substance or substratum) without having any coherent idea of what they are talking about (T 1.4.3.7–8; SBN 222–3).

When he tries to explain personal identity, Hume encounters a similar problem. A certain group of observed causal regularities among perceptions lead us to believe that these perceptions are connected together to form one self. Our belief is justified insofar as these regularities really exist. However, we can find nothing, no connecting principle, that would *explain* why these regularities exist. Those who have tried to provide an account of the connecting principle (linking the perceptions that constitute the self) have run afoul of one or the other of the two principles: either they fail to recognize the distinctness of the distinct parts of the self, or they speak of some metaphysical bonding agent (either a substantial self

in which perceptions inhere or a real connection, like that provided by consciousness when consciousness is supposed to connect the distinct perceptions) without having any coherent idea of what connects the connected parts.

Hume himself speaks of connections among the perceptions that constitute the self, but he makes it clear that his connecting principle does not do what is required of a metaphysical account of the self. As Hume says in the Appendix, summing up his Book I discussion: “If perceptions are distinct existences, they form a whole only by being connected together. But no connexions among distinct existences are ever discoverable by human understanding. We only *feel* a connexion, or a determination of the thought, to pass from one object to another” (T Appendix 20; SBN 635). And this feeling is just one perception among the many that make up the self. It is not capable of binding together distinct perceptions, nor can it explain the observed, and very real regularities we find in the psychological realm. Instead, the feeling itself depends on the prior existence of those regularities.

In the *Appendix*, Hume asserts that the task of explaining “the principles, that unite our successive perceptions in our thought or consciousness” is hopeless. He also says that problems for his account of personal identity arise *only* “when [he] proceed[s] to explain the principle of connexion, which binds [all our particular perceptions] together.” When it comes to that explanation, Hume acknowledges that his account is “very defective” (T Appendix 20; SBN 635). That should not come as news to careful readers, since this “defect” was acknowledged in the original Book I account. The news is that if Hume is right about the implications of his two principles, then this defect is one that *any* coherent and intelligible account of the self must have. Hume’s mention of the only possible solutions: “Did our perceptions either inhere in something . . . individual, or did the mind perceive some real connexion among them” (T Appendix.21; SBN 636) makes it quite clear that the problem he finds with his own account of personal identity is that it does not satisfy the constraint mentioned earlier: that a satisfactory metaphysical account of a composite self should say what directly connects the parts of the self together.

The problem is general. This is the same problem Hume encountered when trying to account for cause and effect and when trying to account for external objects. Here are general grounds for skepticism about metaphysical accounts of any complex phenomena. All these metaphysical explanations require that the mind conceive of a connection between distinct existences. However, if the mind never conceives of connexions among distinct existences, then no satisfactory metaphysical account of any complex thing is possible.

Hume does offer grounds for skepticism about metaphysical accounts of causation, external objects, and the self – accounts that try to *explain* the observed and undeniable regularities that form the basis for our beliefs about these things. Those same grounds, however, do not support skepticism about the existence of causal relations, external objects, and selves, or about our ability to know that when two objects are causally related, which objects are mind-independent and which are

not, and about where my self leaves off and someone else begins. Metaphysical theories of connecting principles would be nice, but they are not required either for the existence of connections or for justified beliefs about them. (See Strawson 1989 for a similar interpretation of the scope of Hume's skepticism.) Hume argues that our experiences of real regularities are the only justification we have for those beliefs, and these experiences provide all the justification we really need, especially since any attempt to go beyond experience in order to provide a metaphysical explanation of experienced regularities leads only to confusion, contradiction, and error – in the intellectual world as well as in the natural.

The fact that the same “contradictions” can be found in philosophical theories of the intellectual world as are found in the natural world would make the case for skepticism about metaphysical theories of *anything* even stronger than Hume had originally thought. That may be why Hume sees his review of the section on personal identity as providing a good *general* reason for skepticism (T Appendix 10; SBN 633). He already had good reasons to doubt that we can have knowledge of the principles that connect together the diverse elements that constitute the system of nature (via cause and effect) and the distinct parts that constitute external objects. Now, when he realizes that the same doubts arise concerning the principles that connect together the perceptions that constitute a mind, the case for skepticism is complete. No connecting principles are knowable, but satisfactory metaphysical accounts of both the mind and the external world require knowledge of connecting principles.

On this reading, Hume's skepticism is general but its target is limited. This skepticism is grounded in the principle that the mind never perceives connections among distinct existences. This is the reason it is limited, extending only to explanations of observed regularities in terms of connecting principles. Once we understand the reasons for skepticism about metaphysical accounts that posit connecting principles, we can see that these reasons need not undermine our beliefs in connections. Such beliefs, which Hume holds are inevitable and indispensable in the conduct of life, can be justified in the absence of satisfactory metaphysical accounts. The justification may be limited, incomplete, and subject to philosophical objections and doubts, but a limited and incomplete justification is better than no justification at all. It is certainly better than a proposed justification that obscures the facts or invents fictions that cannot be maintained consistently. And some philosophical objections and doubts are of such a nature that once raised, they can never be satisfactorily answered. (For example, Descartes's antecedent skepticism, which raises doubts about the reliability of our all epistemic faculties, cannot be answered, since such an answer would require reliance on the very faculties whose reliability has been called into question.)

In assessing Hume's views on personal identity, one question that remains is whether he can account for our beliefs about other persons. While Hume himself does not give such an account, it appears his account can be extended to cover the case of other selves. Although we cannot directly perceive other people's

thoughts and feelings, these thoughts and feelings are made manifest in many ways. Hume notes that they are made manifest in speech, gesture, and action (T 3.2.1.2; SBN 477). If we can perceive our own thoughts, sensations, and feelings and note their external manifestations, and if we can detect causal patterns among both so that we have a sense of self as well as an understanding of the external manifestations and causes of these thoughts and feelings, then as long as we accept the principle that “the same effect never arises but from the same cause” (T 1.3.15, 6; SBN 173), which is one of Hume’s rules for judging cause and effect, we can discover the causal relations that obtain between other people’s perceptions sufficiently to ground judgments concerning other minds.

While much remains to be said about Hume’s theory of personal identity, I close with a brief examination of the practical implications of Hume’s theory of the self. Derek Parfit (1971) has argued that accepting an account of the self that explains selfhood in terms of psychological connections between changing mental contents could effect salutary changes by blurring the boundaries of self-regard and concern for others. While a full discussion of a Humean response to this intriguing suggestion is impossible within the limits of this article, a few sobering remarks seem in order. One of these remarks concerns Hume’s views on the limits of self-construction. For Hume, the self is not a mere illusion, since there really are real regularities among the perceptions that constitute each individual mind. Moreover, the distinctions we make between one self and another are also based on facts about regularities. Whether we are composites of causally connected mental states or singular centers of consciousness, these facts remain as they are and form the basis for the distinctions we make between one self and another. If the facts are undeniable and the synthesizing activities are unavoidable, reflection will do little to undermine our common beliefs about the boundaries of the self and the proper objects of prudential concern.

As regards the therapeutic effects of any philosophical account of the self, Hume held that, generally, and happily for us, very refined reflections have little effect on practice (T 1.4.7.7 and 1.4.7.8; SBN 268). He would likely take this line with respect to metaphysical reflections on the nature of the self. Our prudential concerns are unlikely to be countered by such abstract reflections. If philosophy recommends a widening of concerns beyond the self, it may succeed, but only if something in human nature supports this recommendation. Sympathy is, for Hume, the natural principle in human beings that opens us to the sufferings and happiness of others and gives us a lively concern for them (T 2.1.11.2; SBN 316). But sympathy is based, in part, on our ability to imagine what another person thinks and feels. We often do this by imagining what we would think and feel under similar circumstances (T 2.1.11.5; SBN 317). Sympathy also requires that these imaginings have an affective element that spills over into our thoughts about other people’s feelings. For this to occur, we must first care about our own pains and pleasures, our own happiness and misery. So, here is another paradox for paradox lovers: the borders between self and other must be drawn clearly before

sympathy can soften them. And sympathy, rather than metaphysics, represents our best chance at this self-expansion.

See also 5: BELIEF, PROBABILITY, NORMATIVITY; 6: CAUSATION; 7: IDENTITY, CONTINUED EXISTENCE, AND THE EXTERNAL WORLD.

## References

- Ainslie, Donald (2004) "Hume's anti-cogito." Paper presented at the 31st Annual Hume Conference, Tokyo, Japan.
- Baier, Annette (1979) "Hume on heaps and bundles," *American Philosophical Quarterly* 16, pp. 285–95.
- (1991) *A Progress of Sentiments: Reflections on Hume's Treatise*. Cambridge, MA: Harvard University Press.
- Capaldi, Nicholas (1989) *Hume's Place in Moral Philosophy*. New York: Peter Lang.
- Chazan, Pauline (1992) "Pride, virtue, and self-hood: a reconstruction of Hume," *Canadian Journal of Philosophy* 22, pp. 45–64.
- Fieser, James (1989) "Hume's pyrrhonianism: a developmental interpretation," *Hume Studies* 15, pp. 93–119.
- Fogelin, Robert J. (1985) *Hume's Skepticism in the Treatise of Human Nature*. London: Routledge.
- Garrett, Don (1981) "Hume's self-doubts about personal identity," *Philosophical Review* 90, pp. 337–58.
- (1997) *Cognition and Commitment in Hume's Philosophy*. New York: Oxford University Press.
- Henderson, Robert S. (1990) "David Hume on personal identity and the indirect passions," *Hume Studies* 16, pp. 33–44.
- Kemp Smith, Norman (1941) *The Philosophy of David Hume*. London: Macmillan.
- Locke, John (1975) *An Essay concerning Human Understanding*, ed. P. H. Nidditch. Oxford: Clarendon Press.
- McIntyre, Jane L. (1989) "Personal identity and the passions," *Journal of the History of Philosophy* 27, pp. 545–57.
- (1993) "Norms for a reflective naturalist: a review of Annette C. Baier's *A Progress of Sentiments*," *Hume Studies* 19, pp. 317–23.
- Noonan, Harold W. (2003) *Personal Identity*, 2nd edn. London and New York: Routledge.
- Olson, Eric T. (1997) *The Human Animal: Personal Identity without Psychology*. New York: Oxford University Press.
- Parfit, Derek (1971) *Reasons and Persons*. Oxford: Clarendon Press.
- Patten, S. C. (1976) "Hume's bundles, self-consciousness and Kant," *Hume Studies* 2, pp. 59–75.
- Pears, David (1990) *Hume's System: an Examination of the First Book of his Treatise*. Oxford and New York: Oxford University Press.
- Pike, Nelson (1967) "Hume's bundle theory of the self: a limited defense," *American Philosophical Quarterly* 4, pp. 159–65.

- Purviance, Susan M. (1997) "The moral self and the indirect passions," *Hume Studies* 23, pp. 195–212.
- Robison, Wade L. (1974) "Hume on personal identity," *Journal of the History of Philosophy* 12, pp. 181–93.
- Strawson, Galen (1989) *The Secret Connexion: Causation, Realism, and David Hume*. Oxford: Clarendon Press.
- Stroud, Barry (1977) *Hume*. London: Routledge.
- Swain, Corliss (1991) "Being sure of one's self: Hume on personal identity," *Hume Studies* 17, pp. 107–24.
- Swoyer, Chris (1982) "Hume and the three views of the self," *Hume Studies* 8, pp. 43–61.
- Traiger, Saul (1985) "Hume on finding an impression of the self," *Hume Studies* 11, pp. 47–68.
- (1987) "Impressions, ideas, and fictions," *Hume Studies* 13, pp. 381–99.
- Waxman, Wayne (1992) "Hume's quandary concerning personal identity," *Hume Studies* 18, pp. 233–54.
- Winkler, Kenneth P. (2000) "'All is revolution in us': personal identity in Shaftesbury and Hume," *Hume Studies* 26, pp. 33–40.

### Further reading

- Ainslie, Donald (2001) "Hume's reflections on the simplicity and identity of mind," *Philosophy and Phenomenological Research* 62, pp. 557–78.
- Ashley, L., and M. Stack (1974) "Hume's theory of the self and its identity," *Dialogue* 13, pp. 239–54.
- Baxter, Donald L. M. (1998) "Hume's labyrinth concerning the idea of personal identity," *Hume Studies* 24, pp. 203–34.
- Beauchamp, Tom L. (1979) "Self-inconsistency or mere perplexity?," *Hume Studies* 5, pp. 37–44.
- Biro, John I. (1976) "Hume on self-identity and memory," *Review of Metaphysics* 30, pp. 19–38.
- (1979) "Hume's difficulties with the self," *Hume Studies* 5, pp. 45–54.
- Bricke, John (1977) "Hume on self-identity, memory and causality," in G. P. Morice (ed.) *David Hume: Bicentenary Papers* (pp. 167–74). Austin: University of Texas Press.
- (1980) *Hume's Philosophy of Mind*. Princeton, NJ: Princeton University Press.
- Chisholm, Roderick (1976) *Person and Object*. LaSalle, IL: Open Court.
- Cummins, Phillip D. (1973) "Hume's disavowal of the *Treatise*," *Philosophical Review* 82, pp. 371–9.
- Flage, Daniel E. (1990) *David Hume's Theory of Mind*. London and New York: Routledge.
- Giles, James (1993) "The no-self theory: Hume, Buddhism, and personal identity," *Philosophy East and West* 43, pp. 175–200.
- Immerwahr, John (1977) "The failure of Hume's *Treatise*," *Hume Studies* 3, pp. 57–71.
- Loeb, Louis E. (1992) "Causation, extrinsic relations, and Hume's second thoughts about personal identity," *Hume Studies* 18, pp. 219–32.

- MacNabb, D. G. C. (1951) *David Hume*. London: Hutchinson University Library.
- Martin, Raymond, and John Barresi (2000) *Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century*. London: Routledge.
- McIntyre, Jane L. (1979a) "Further remarks on the consistency of Hume's account of the self," *Hume Studies* 5, pp. 55–61.
- (1979b) "Hume on the self," in D. F. Norton, N. Capaldi, and W. L. Robison (eds.) *McGill Hume Studies* (pp. 79–88). San Diego: Austin Hill Press.
- (1993) "Hume's underground self," *Studies in Early Modern Philosophy* 3, pp. 110–26.
- (2003) "'So great a question': a critical study of Raymond Martin and John Barresi, *Naturalization of the Soul: Self and Personal Identity in the Eighteenth Century*," *Hume Studies* 29, pp. 363–73.
- Nathanson, Stephen (1976) "Hume's second thoughts on the self," *Hume Studies* 2, pp. 36–45.
- Noxon, James (1969) "Senses of identity in Hume's *Treatise*," *Dialogue* 8, pp. 367–84.
- (1973) *Hume's Philosophical Development*. Oxford and New York: Clarendon Press.
- Passmore, John (1952) *Hume's Intentions*. Cambridge: Cambridge University Press.
- Pears, David F. (1966) "Hume on personal identity," in D. F. Pears (ed.) *David Hume: a Symposium* (pp. 43–54). New York: St Martin's Press.
- (1993) "Hume on personal identity," *Hume Studies* 19, pp. 289–300.
- Penelhum, Terence (1955) "Hume on personal identity," *Philosophical Review* 54, pp. 571–89.
- (1967) "Personal identity," in P. Edwards (ed.) *Encyclopedia of Philosophy*, vol. VI (pp. 95–107). New York: Macmillan.
- (1975a) *Hume* (esp. ch. 4). London: Macmillan.
- (1975b) "Hume's theory of the self re-visited," *Dialogue* 14, pp. 389–409.
- Pitson, A. E. (2002) *Hume's Philosophy of the Self*. London: Routledge.
- Reid, Thomas (1892) *The Philosophy of Reid, as contained in the Inquiry into the Human Mind on the Principles of Common Sense*, ed. E. H. Sneath. New York: Henry Holt.
- (1969) [1785] *Essays on the Intellectual Powers of Man*. Cambridge, MA: MIT Press.
- Robison, Wade L. (1979) "In defense of Hume's *Appendix*," in D. F. Norton, N. Capaldi, and W. L. Robison (eds.) *McGill Hume Studies* (pp. 89–99). San Diego: Austin Hill Press.
- Roth, Abraham Sesshu (2000) "What was Hume's problem with personal identity?" *Philosophy and Phenomenological Research* 61, pp. 91–114.
- Yang, Xiaomei (1998) "Hume's view on personal identity: scepticism or nonscepticism?" *History of Philosophy Quarterly* 15, pp. 37–56.