

# Hume's Theory of Ideas

DON GARRETT

Thomas Reid, David Hume's philosophical contemporary and fellow Scot, characterized many early modern philosophers, from Descartes to Hume, as holding "the common theory of ideas." By this he meant that they regarded the mind as *immediately* perceiving only certain mental entities, usually called *ideas*. Although Hume differed from most of his early-modern predecessors in using the term "perception" for these mental entities, reserving the term "idea" for the proper subset of them experienced in thought as opposed to feeling, Reid was clearly right to classify Hume as part of this tradition; for Hume confidently asserts, "'twill readily be allow'd that . . . nothing is ever really present to the mind, besides its own perceptions" (T 1.4.2.21). Yet the label Reid helped to popularize can easily obscure many important differences among the philosophers to whom he applied it, and this is nowhere truer than in the case of Hume.

Hume's philosophical ambition, as expressed in the Introduction to *A Treatise of Human Nature*, was to establish a "science of man" that, by explaining the operations of the human mind, would provide a foundation "almost entirely new" for all of the sciences. The primary objects of that foundational science are the mental entities that he calls perceptions, and especially those perceptions that he classifies as ideas. There is a way, then, in which most of his philosophy is a "theory of ideas" even in his narrower sense of the term "idea," and a fortiori in Reid's broader sense. This essay, however, will be limited to considering (1) the most basic distinctions he draws among kinds of perceptions and (2) the most basic principles that he propounds concerning their operations. As one would expect, grasping these is essential to understanding Hume's philosophy as a whole. In examining these distinctions and principles, we will have occasion to note some of the uses to which he puts them, some of the objections that have been raised against them, and some of the important ways in which Hume's version of "the common theory of ideas" differs from those of his predecessors.

## Basic Distinctions

Although Hume briefly recapitulates many basic aspects of his theory of ideas in *An Enquiry concerning Human Understanding*, his fullest account of it by far is to be found in *A Treatise of Human Nature* – primarily, although not exclusively, in Book 1, part 1

("Of ideas, their origin, composition, connexion, abstraction, &c."). He begins that work by drawing two basic distinctions between kinds of perceptions generally. He then goes on to draw several basic distinctions within the domain of those perceptions he calls *impressions* and several basic distinctions within the domain of those perceptions he calls *ideas*. A final important distinction between kinds of perceptions is not drawn as explicitly but emerges gradually in the course of the text.

### *Distinctions between kinds of perceptions*

The opening paragraph of the main body of the *Treatise* is devoted to the distinction between *impressions* and *ideas* itself: "The difference betwixt these consists in the degrees of force and liveliness, with which they strike upon the mind, and make their way into our thought or consciousness" (T 1.1.1.1). Impressions include "all our sensations, passions and emotions, as they make their first appearance in the soul"; ideas include "the faint images of these in thinking and reasoning," such as "all the perceptions excited by the present discourse, excepting only, those which arise from the sight and touch, and excepting the immediate pleasure or uneasiness it may occasion." The difference between having impressions and having ideas is thus that "between feeling and thinking."

In addition to "force and liveliness," Hume frequently uses the term "force and vivacity," or simply "liveliness" or "vivacity," to designate the difference between impressions and ideas. The term "force" suggests a degree of causal efficacy, while the terms "liveliness" and "vivacity" suggest a more purely phenomenal difference. Since it proves to be a central tenet of Hume's philosophy that the causal efficacy of a quality can never be phenomenally present in the quality itself, these characterizations are not obviously equivalent. Evidently, Hume's view is that impressions and ideas differ in their degrees of a particular phenomenal quality, degrees that are reliably correlated with differences in causal efficacy. The terms "liveliness" and "vivacity" naturally suggest degrees of an ordinary kind of perceptible character, such as visual brightness or auditory loudness, and Hume himself suggests degrees of brightness (which can vary while the shade of color remains the same) as an analogy (T 1.3.7.5); but the analogy can mislead. For he requires that degrees of liveliness or vivacity (which apply to all perceptions and not merely to those of one sense modality) be degrees of a distinctive kind of phenomenal "manner" that does not alter the qualitative character of the perception itself, so that an impression and an idea can be identical in qualitative character even in such respects as brightness or loudness while differing *only* in this one further dimension.

Hume concludes the first paragraph of the *Treatise* by remarking that, despite the ease with which the distinction can usually be drawn, "our ideas may approach to our impressions" in sleep, fever, or madness, and that, on the other hand, our impressions "can become so faint and low, that we cannot distinguish them from our ideas." This concluding remark has seemed to many readers to be incompatible with the opening of the paragraph, for it seems to imply that some impressions and ideas do *not* differ in force and vivacity. Closer examination reveals that the remark does not quite entail that conclusion; for ideas may "approach" impressions in vivacity without ever reaching them, and a difference may exist that the mind cannot reliably distinguish (since

distinguishing can also require accurate memory and skill at fine comparison). Still, it is clear that Hume is supposing that impressions and ideas typically have different kinds of causes as well as different kinds of effects, and he is at least as concerned with these as he is with their intrinsic phenomenal character.

The second basic distinction that Hume draws between kinds of perceptions is that between *simple* and *complex* perceptions:

There is another division of our perceptions, which it will be convenient to observe, and which extends itself both to our impressions and ideas. This division is into SIMPLE and COMPLEX. Simple perceptions or impressions and ideas are such as admit of no distinction nor separation. The complex are the contrary to these, and may be distinguished into parts. Tho' a particular colour, taste, and smell are qualities all united together in this apple, 'tis easy to perceive they are not the same, but are at least distinguishable from each other. (T 1.1.1.2)

It has often been assumed that Hume's distinction between simple and complex perceptions corresponds exactly to John Locke's distinction between simple and complex ideas. In *An Essay concerning Human Understanding* (Locke 1690), Locke identifies as simple the ideas of particular colors, sounds, tastes, and smells, as well as heat and cold; solidity, extension (i.e., the having of extent in spatial dimensions), figure, space, motion, and rest; volition and perception; and pleasure, pain, unity, succession, power, and existence. In some cases, he holds, there are necessary connections among such simple ideas, and these can provide an important source of knowledge. Hume, however, regards many of these ideas as complex – including those of space, extension, figure, motion, and succession – because he regards the spatial or temporal complexity of an idea as incompatible with its simplicity. (The reference to the color of the apple in his example concerns a *type* of simple idea, since on his view a perception of an apple must be composed, in part, of a large number of colored but spatially indivisible *minima*, and each of these is a perception of color in its own right.) Furthermore, he denies that there is any separate or distinct simple perception of unity or existence. Hume's way of drawing the simple/complex distinction, together with his theory of abstract ideas, allows him to deny that there are any necessary connections between different simple ideas.

### *Distinctions between kinds of impressions*

Every impression, Hume holds, is either an *impression of sensation* or an *impression of reflection*. The former, he writes, arise “in the soul originally, from unknown causes,” while the latter arise “in great measure” from ideas that have been copied in memory or imagination from impressions of sensation (T 1.1.2.1). Impressions of reflection thus include all of the passions, such as love, hatred, pride, humility, anger, benevolence, hope, fear, and desire. They also include sentiments, such as those of approbation and disapprobation, that are more delicate than the passions, as well as other feelings that arise in the mind in the course of its operations with ideas, such as impressions of “determination” or “necessity” and “facility” or “ease.” Hume's claim that impressions of sensation “arise . . . from unknown causes” should not be taken as an expression of total skepticism about the existence of external objects of sense perception. One reason

for caution on this score is that he may be intending only to express ignorance of the particular means by which external objects, operating through complex sense organs, nerves, and brain structures, produce impressions in the mind. But even if his denial of knowledge is meant to extend to the very existence of external objects as the causes of impressions of sensation, his official standards for *knowledge* are extremely high, encompassing mathematics and a few other intuitive or demonstrable truths but no causal relations, so that to say that one lacks “knowledge” of something in this sense is not to deny that one can have a strongly warranted belief about it. On the contrary, he has already indicated with a high degree of confidence, just a few paragraphs earlier, that we typically have sensory impressions as the result of the stimulation of sense organs by external objects. Later in the *Treatise*, he states without hesitation that impressions of sensation arise “from the constitution of the body, from the animal spirits, or from the application of objects to the external organs” (T 2.1.1.1).

Within the domain of impressions of sensation, Hume acknowledges the common distinction between those of *primary qualities* (such as solidity, extension, shape, size, and motion) and *secondary qualities* (such as colors, sounds, tastes, smells, heat, and cold) of bodies. This distinction is of course derived from Locke, who defines *qualities* of bodies as “powers to produce ideas” in minds and asserts that the primary qualities do, while the secondary qualities do not, *resemble* the ideas (i.e., Humean impressions) that they produce there (*Essay* II.viii). Hume does not unambiguously endorse this doctrine of what he calls “the modern philosophy” concerning secondary qualities, but he does allow that the modern philosophers have one “satisfactory” causal argument for it (T 1.4.4, “Of the Modern Philosophy”), derived from the relativity of their perception (see chapter 7).

The manner in which qualities of bodies *can* resemble impressions is quite different, however, in Hume from what it is in Locke. For Locke – as for Descartes and indeed nearly all of the early modern philosophers except Spinoza and Malebranche – the human mind is a substance, of which ideas are mere qualities or *modes*. Thus, Locke specifically compares the relation in which minds stand to their ideas with the relation in which bodies stand to their motions. As mere modes of a substance that, on Locke’s view, is probably immaterial and unextended, “ideas” of primary qualities are not literally extended, solid, figured, or sized. So how can he suppose that they *resemble* these qualities of body? Presumably, Locke is thinking of the ideas of primary qualities as containing those qualities through an intrinsic representational capacity – containing them *objectively*, as Descartes puts it, rather than *formally*. Descartes’ use of the term “objective” in connection with this intrinsic representational capacity derives from the notion of the *objective being* of a thing, originally conceived by Descartes’ scholastic forebears as one of two ways of being that a thing could have; thus, a thing could have formal being in reality, objective being in an idea in the mind, or both.

Hume, in contrast, recognizes no such distinction between formal and objective being, even implicitly. On the contrary, he insists that there is only one kind of existence (T 1.2.6, “Of the idea of existence and of external existence”). For him, perceptions of *extension* are literally extended (and hence complex rather than simple, as we have already noted); perceptions of *squares* are literally square; and perceptions of *red* are literally red. The question of whether the objects causing these impressions *have qualities that resemble them* is therefore a quite straightforward one, even if it is difficult in some cases

to establish the correct answer. At the same time, however, this Humean literalism about the qualities of perceptions makes it correspondingly more difficult to think of perceptions as *modes of a thinking substance*, and indeed Hume denies that they are; instead, he holds, the mind is a collection or “bundle” of causally interrelated mental particulars, some extended and others (such as smells and passions) unextended:

the true idea of the human mind is to consider it as a system of different perceptions or different existences, which are linked together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other. (T 1.4.6.19; see ch. 8)

### *Distinctions between kinds of ideas*

Because ideas share the qualitative characters of impressions while differing in force and vivacity, any distinction originally drawn within the domain of impressions can readily be matched by a corresponding distinction within the domain of ideas (e.g., between ideas of impressions of sensation and ideas of impressions of reflection, or between ideas of primary qualities and ideas of secondary qualities). However, Hume's most general distinction specifically within the domain of ideas is that between ideas of *memory* and ideas of *the imagination*. These differ in two main respects. First, ideas of memory have more force and liveliness than ideas of the imagination, although not as much force and liveliness as impressions. Second, while the mind may voluntarily alter the order or arrangement of ideas of imagination, ideas of memory preserve a certain fixed order that is not subject to voluntary control. (The distinction between memory and imagination is explored extensively in chapter 3.)

Within the domain of ideas of imagination, Hume draws a number of further distinctions. One of these is the distinction between *beliefs* and *mere ideas* (i.e., those that are not beliefs) – which is again a distinction of degrees of force and liveliness, with beliefs having less force and liveliness than memories but more than is possessed by those ideas that are not beliefs. The processes by which ideas acquire the force-and-liveliness that makes them beliefs is a topic explored at length in chapter 6; but it is worth emphasizing that for Hume belief in its most basic form is not an act of judgment that employs concepts, but is rather the having of a particular mental image – i.e., a mental map or model – that, in virtue of its liveliness, is able to motivate and guide behavior.

Although Humean belief need not involve concepts, many of them nonetheless do. What we would call *concepts*, Hume calls *abstract ideas*. For Locke, abstract ideas achieve generality through intrinsic indeterminacy. The abstract idea of a man, for example, would simply be inherently indeterminate with respect to height or color, while the abstract idea of a triangle would be inherently determinate with respect to the number of angles but inherently indeterminate with respect to their size; similarly, the idea of unity would represent oneness that was not the oneness of any particular thing. Because Hume implicitly but firmly rejects the notion of ideas representing through their containing “objectively” that which they represent, he cannot accept the Lockean account of generality; an inherently indeterminate idea would have to be literally (i.e., “formally”) indeterminate, and that would be no more possible than the existence of an indeterminate object. Instead, he agrees with George Berkeley that all

ideas are fully determinate in their own nature and achieve generality only through their use. On Hume's account, an idea becomes an abstract idea when it is associated with a "general term" of a language in such a way that the particular idea (which we may call the *exemplar*) is disposed to call up other resembling ideas (which we may call the *revival set*) as needed in thinking or reasoning (T 1.1.7, "Of abstract ideas"). Thus the abstract idea (i.e., the exemplar) of TRIANGLE will be the idea of a triangle determinate in angles, size, and color; it may be red and equilateral, for example. Yet if a claim is made that "all triangles are equilateral," an idea of a right triangle or of an obtuse triangle will come readily to mind, leading to the denial of the claim. To form the conceptual judgment that something exists as a triangle is presumably to include a lively idea of it within the revival set of the abstract idea of TRIANGLE.

It is sometimes objected to Hume's theory of abstract ideas that the activity of coming to associate a variety of resembling particulars with an exemplar under a common term requires that one *already* possess the concept of the respect in which they resemble one another (e.g., Kemp Smith 1941: 257; Mounce 1999: 27–8; Johnson 1995: 73–5). This objection, however, misses the point of his naturalistic theory of concept acquisition, which requires only that similar instances be able to produce similar effects in the mind, including the idea of a similar word. For Hume, similar causes can often produce similar effects, whether outside the mind or inside it, prior to the existence of any *concept* of the respect of similarity.

Humean abstract ideas thus differ from ideas of particulars not through their intrinsic character but through their causal and functional relations. The distinction between ideas of *substances* and ideas of *modes* is another distinction of this kind. Ideas of both kinds are complex ideas composed of ideas of particular qualities and designated by a single term. In the case of ideas of substances, however, additional ideas may be added to the complex without changing the term, whereas any addition to the complex idea of a mode requires a new term (T 1.1.6, "Of modes and substances").

For Hume, thinking about something requires having an idea of it. As is therefore implied by his discussion in the *Treatise* of external objects ("bodies"), impressions, and ideas, he recognizes ideas (of memory and imagination) of external objects, of impressions, and of ideas themselves. Indeed, since he also discusses ideas of ideas, it follows that he even recognizes ideas of ideas of ideas. Yet how can qualitatively similar ideas represent such different things? The answer must lie in his theory of representation.

### *Representation: a final distinction between kinds of perceptions*

All ideas are *representations*, according to Hume, for he remarks that "ideas always represent their objects or impressions" (T 1.3.14.6; see also T 1.2.3.11). In contrast, it is often supposed that, on his view, impressions cannot represent anything at all (see, for example, Owen and Cohon 1997). In fact, however, while he straightforwardly denies that the passions "contain any representative quality" or have any "reference to any other object" (T 2.3.3.5), he never denies that impressions of sensation do. On the contrary, he regularly writes of impressions "of" various external objects. He also claims explicitly that some impressions of sensation "represent" extension or "represent" large external objects as being minute (T 1.2.1.5 and T 1.2.3.15). And in *An Enquiry concerning Human Understanding*, he remarks concerning impressions of sensation:

[N]o man, who reflects, ever doubted, that the existences, which we consider, when we say, this house and that tree, are nothing but perceptions in the mind, and fleeting copies or *representations* of other existences, which remain uniform and independent. (EHU 12.9, emphasis added)

The interpretation of Hume as denying that impressions of sensation can represent rests largely on the following remark from the *Treatise*:

Now since nothing is ever present to the mind but perceptions, and since all ideas are deriv'd from something antecedently present to the mind; it follows, that 'tis impossible for us so much as to conceive or form an idea of any thing specifically different from ideas and impressions. (T 1.2.6.8)

If impressions of sensation represent anything, it seems, they represent bodies and their qualities; but it also seems that whatever is inconceivable cannot be represented by perceptions in the mind. Hence, if the mind cannot conceive anything “specifically different” from perceptions, one might reason, impressions of sensation cannot represent bodies or their qualities, and so cannot represent at all.

This interpretation, however, depends on a misunderstanding of what Hume means by “specifically different from ideas and impressions.” He is not denying that we can conceive of bodies, understood as objects having what he calls a “continu'd and distinct existence” – that is, as (1) continuing to exist when not perceived by the mind and (2) having an existence that is distinct from the mind in having an external location and causally independent existence and operation. Indeed, a later section of the *Treatise* (T 1.4.2, “Of skepticism with regard to the senses”) is largely devoted to explaining how we *do* conceive objects in this way. Rather, he is denying only that we can conceive of such bodies either as having a *separate species of existence* (e.g., Lockean “real existence” understood as just one kind of existence in contrast with ideal or “objective” existence) or as having specific, conceivable intrinsic qualities (e.g., color, sound, taste, smell, heat, cold, shape, motion) different from those of perceptions. Hume immediately goes on to explain our only recourse for conceiving of bodies as having intrinsic qualities different from these:

The farthest we can go towards a conception of external objects, when suppos'd *specifically* different from our perceptions, is to form a relative idea of them, without pretending to comprehend the related objects. (T 1.2.6.9)

That is, we can conceive of the qualities of bodies in a relative way merely as *whatever is the cause of the qualities of our perceptions*. This is in no way incoherent – Hume recognizes the conceptual legitimacy of “relative ideas” by which one conceives of an unknown relatum by conceiving it *as* what stands in a known relation to a known relatum – but it does not capture any “specific” difference, since the unknown qualities of bodies that would cause the qualities of perceptions are left unspecified. “Generally speaking” however, he continues, “we do *not* suppose them [i.e., bodies] specifically different [even in this way]; but *only* attribute to them different relations, connexions, and durations” (emphasis added) – referring the reader in a footnote to *Treatise* 1.4.2, where the conception of bodies as continued and distinct existences

through the attribution of such different “relations, connexions, and durations” is fully explained.

Thus, Hume does not deny that impressions of sensation represent. Indeed, more broadly, he nowhere denies that impressions of reflection *other than the passions* represent. On the contrary, he seems to think of moral and aesthetic sentiments as representing the moral and aesthetic qualities or persons or objects (virtue and vice, beauty and deformity, respectively) for which they constitute a kind of “sense.” Similarly, when he writes of impressions “of” mental determination or “of” facility in mental transition, he seems to think of these impressions as representing the features of the mind’s operations that they reliably serve to indicate.

Exactly why some but not all perceptions represent is something that Hume does not explicitly explain. However, he does provide enough clues to allow readers to come to a reasonable conclusion about the answer. First, he makes it clear that no perceptions – not even ideas – are *inherently* representational; rather, “the reference of the idea to an object [is] an extraneous denomination, of which in itself it bears no mark or character” (T 1.1.7.6). In adopting this doctrine, he is departing considerably and crucially from his most influential predecessors in “the common way of ideas” who, as we have seen, regard at least many ideas as representing what they do in virtue of an intrinsic representational property by which they contain the objective reality of that which they represent. Second, he makes it clear early in the *Treatise* that the resemblance of ideas to the impressions from which they are causally derived, in a way that amounts to *copying*, plays a fundamental role in their ability to represent them. Yet ideas can represent without *entirely* resembling what they represent, on Hume’s view, for his definition of “falsehood” (T 3.1.1.9; see also T 2.3.10.2) requires that false ideas represent things as having qualities or relations that they do not have. More importantly, resemblance and causal derivation, even when together constituting a kind of copying, are not *in general* sufficient for representation (for example, decorative motifs in a building’s architecture may be copied from those of another building without representing them), and he could reasonably be expected to recognize that fact. Third, Hume gives a number of examples of *non-mental* representation throughout his works: words can represent “objects and facts” and “sentiments and impressions”; the space of a theater can represent the place where a play is set; children can represent their parent’s family (in the eliciting of passions); money can represent the beautiful and agreeable objects that it affords the power of obtaining; the giving of “stone and earth” can represent the conveyance of a manor in property law; and a “taper, habit, or grimace” may represent a religious mystery. In each of these cases, he writes of *representation* precisely because the representing item or event is taking on a significant part of the causal and/or functional role of what is said to be represented, and it is doing so through its production of mental effects such as ideas, beliefs, sentiments, passions, and volitions – or at least dispositions to them. In some cases, the representation plays the *causal* role of the represented object by reliably producing its typical mental effects and dispositions as a causal intermediary. Such representation occurs, for example, when a word is used to indicate the presence of the thing named – or when an impression of sensation indicates the presence of a body, which in turn produces effects in the mind through the mediation of that impression. This is representation through *indication*. In some cases, the representation plays the *functional* role of the represented object by

producing mental effects and dispositions that *replicate* or *parallel* effects and dispositions typically produced by the represented object itself in similar or parallel circumstances. This occurs, for example, when the space of a theatrical stage contributes to the production in the minds of audience members of the same mental effects or dispositions that would be produced in the minds of auditors of the events represented – or when a *belief* in the existence of some object produces mental effects and dispositions parallel to those that would be produced in the external world and/or the mind by the presence of the object itself. Similarly, even a *mere idea* of an object will typically give rise at least to *ideas of its typical mental effects*. This is representation through *modeling*. Both kinds of representation – indication and modeling – may be combined in a single case, as when an object causes an impression of sensation that thereby indicates the object causing it, while also giving rise to beliefs about the object's causal consequences that parallel the actual consequences of that object in the external world, thereby modeling it as well.

If this is indeed Hume's understanding of representation in general, it would explain why he regards both resemblance and causal derivation as particularly *conducive* to representation, even if (as is evident from some of his examples of non-mental representation) they are not always required. For resembling objects are *more likely* to have resembling effects, particularly on the mind. And causal derivation, in addition to being essential to *indication* (i.e., playing the causal role of something else as an intermediary), also often serves to determine more precisely *what* object's functional role is being played in the production of mental effects and dispositions. *Mental representation* – that is, representation by perceptions – would then be a special case of representation in general for Hume: specifically, it would be the special case in which not only the effects and dispositions produced but also the representations themselves were mental entities. Representation *by ideas* would then be a yet further special case. Because ideas have the causal roles they do almost exclusively through their phenomenal properties, causal roles that naturally mimic those of the perceptions that they resemble and from which they are causally derived, it is understandable that copying should be essential or nearly essential to representation, as Hume indicates it is, in the special case of ideas.

That this is Hume's general understanding of mental representation is confirmed by the fact that differences in the representational capacities of individual ideas that are identical in phenomenal character – such differences, for example as representing all members of a general class as opposed to just a particular, or representing a substance as opposed to a mode – are determined by the causal and functional role of the ideas in question. At the same time, a causal/function understanding of representation would serve to explain how one idea represents only an impression, while another idea identical in phenomenal character represents a body, or even another idea. Thus, an idea will represent an *impression* by mediately producing some of the effects of that impression and also by modeling it, producing mental effects and dispositions that parallel the effects that the impression itself produces. An idea will represent a *body* when it also produces *additional* effects and dispositions – namely, effects and dispositions parallel to those that the impression would have produced if it had been present at a time at which it was not actually perceived. For in this way, the idea represents its object as having a continued and distinct existence. Ideas *of ideas* will differ from ideas of impressions or bodies in two respects. First, they will tend to produce only *ideas of*

passions, sentiments, and volitions. Second, and more importantly, they will be accompanied by an idea *of* the distinctive feeling of the mind that, according to Hume, always accompanies ideas:

[W]e need not be surpriz'd to hear of the remembrance of an idea; that is, of the idea of an idea, and of its force and vivacity superior to the loose conceptions of the imagination. In thinking of our past thoughts we not only delineate out the objects of which we were thinking, but also conceive the action of the mind in the meditation, that certain *je-ne-scai-quoi*, of which it is impossible to give any definition or description, but which every one sufficiently understands. (T 1.3.8.16)

Ideas *of ideas of ideas*, we may suppose, will tend to generate *ideas of ideas* of sentiments, passions, and volitions, and will be associated with a *pair* of ideas of the distinctive feeling of mind that always accompanies ideas: one corresponding to the occurrence of the original idea and one corresponding to the occurrence of the idea of the idea. (For further discussion of Hume's theory of representation, see Garrett 2006.)

## Basic Principles

One notable distinction common to many early modern philosophers that is *not* to be found in Hume is the distinction between *ideas of the intellect* and *ideas of the imagination*. Descartes, Spinoza, and Leibniz all recognize two classes of ideas: (1) radically non-imagistic ideas whose content is innate and (2) imagistic ideas (understanding the term "imagistic" broadly enough to include auditory, gustatory, olfactory, and even emotional images) whose content is derived from experience. Hume, however, like Locke and Berkeley before him, recognizes no ideas of the former, intellectual, kind. In effect, then, he aims to explain all human cognition with only the two representational faculties of imagination and memory, both understood as providing imagistic representations whose content is derived from experience. He invokes four basic principles in the course of his efforts to do so.

### *The Copy Principle*

Standing in a particularly intimate relation to Hume's endeavor to avoid ideas of Cartesian intellect is the most basic and best-known principle of his science of man, which has come to be called the *Copy Principle*: "That all our simple ideas in their first appearance, are derived from simple impressions, which are correspondent to them, and which they exactly represent" (T 1.1.1.7). These simple ideas may then be used to construct new complex ideas – for example, of a unicorn or the New Jerusalem – to which no previous complex impression need have corresponded.

Critics have sometimes accused Hume of treating the Copy Principle as a priori (see, for example, Flew 1961: 22–3; Bennett 1971: 226–7). This would be an inconsistency on his part since, as a principle about causal dependence (specifically, of ideas on prior impressions), it is a proposition of the kind that he repeatedly insists can be established only by experience. In Hume's defense, however, it may be observed that he tries to establish the principle with two appropriately empirical arguments: first, when one has

had both a simple idea and a simple impression with the same character, one finds that the simple impression has always preceded the first appearance of the idea in one's experience; and, second, when one is unable to have a particular simple impression, either through a defect in a sense organ or some other lack of opportunity, one also lacks the corresponding simple idea. Hume characterizes the Copy Principle as a clearer and more precise version of Locke's famous denial of innate ideas, a denial that Locke aims to confirm by giving experiential derivations for all of the crucial intellectual ideas that Descartes had characterized as innate; and, indeed, Hume could cite any of the Lockean attempts to provide such derivations that he regards as successful as partial additional confirmations of the Copy Principle.

One common source of consternation with Hume's treatment of the Copy Principle is his seemingly cavalier dismissal of the notorious "missing shade of blue." As he describes the case, a series of closely resembling shades of blue are arranged from lightest to darkest, with one shade removed from the series. Even without conducting the experiment, he readily grants – as he expects his readers will also – that an observer would be able to imagine an idea of the missing shade without having had an impression of that particular shade before. While admitting that such an occurrence would be an "exception" to his principle, he concludes that "the instance is so particular and singular, that it is scarce worth our observing, and does not merit that, for it alone, we should alter our general maxim" (T 1.1.10).

Although strictly speaking the instance is not entirely singular – similar examples could be constructed with missing tones, or even missing degrees of heat – such counter-examples would (even if they were actually to occur) constitute "near misses" for the Copy Principle, inasmuch as the character of the idea of the missing shade of blue (and its analogs in the other cases) is clearly very closely derived, if not quite precisely copied, from the impressions presented. For this reason, the instances do not constitute serious impediments to the two uses to which Hume actually puts the Copy Principle. First, it underwrites a methodological directive: where one is uncertain of the nature of an idea, try to trace it to the impression from which it is copied, because the impression will have additional force and liveliness, making it less subject to obscurity and confusion. This directive is invoked most explicitly in the case of the idea of necessary connection. Second, the principle provides confirmatory evidence against the existence of purported ideas whose existence is already questionable on other grounds:

When we entertain, therefore, any suspicion, that a philosophical term is employed without any meaning or idea (as is but too frequent), we need but enquire, *from what impression is that supposed idea derived?* And if it be impossible to assign any, this will serve to confirm our suspicion. (EHU 2.9)

The suspect terms to which Hume applies this principle are relatively few: "substance" (as intended to designate a subject in which modes, qualities, or perceptions inhere), "vacuum" (intended as existing space that is empty), "eventless time" – and "power" and "necessary connexion" when these two latter term are intended to designate a genuinely necessitating capacity intrinsic to some cause or a genuinely necessitating relation located in the cause and effect themselves. But in each of these cases, he also has antecedent grounds to be suspicious of the meaningfulness of the term in

question. In the case of “substance,” for example, the idea of substance would have to be of something that could bestow simplicity at a time and lack of change through time on things that are already seen to be complex and changing. Similarly, the perception of a necessary connection in nature would “amount to a demonstration” of the inseparability of two distinct events, the separation of which can already be seen to be possible. In no case is the Copy Principle his sole reason for pronouncing a philosophical term meaningless; the sense that Hume’s uses of the Copy Principle treat it as a priori derives in part from failure to appreciate this point.

Some critics (e.g., Kemp Smith 1941: ch. 14; Johnson 1995: 90–3) have argued that the missing shade of blue is not the only exception that Hume allows to the Copy Principle, for he allows that we have ideas of space, time, and existence even while denying that we have any “separate,” “distinguishable,” or “particular” impression of them. It would be odd indeed if the idea of space were a counterexample to the Copy Principle, for he emphasizes the importance of the principle in the midst of his explanation of the nature of that idea (T 1.2.3.1). But in fact, his account of abstract ideas makes it clear why the Copy Principle does not require an impression of space, time, or existence that is separable from the impressions of the things that are in space, time, or existence. For the ideas of space, time, and existence, as abstract ideas, each consists of a particular idea, serving as an exemplar, associated with a general term in such a way as to be disposed to elicit ideas of a revival set. The exemplar of the abstract idea of space must be a complex idea whose parts are in a spatial arrangement, and so must each of the ideas in its revival set. Similarly, the exemplar of the abstract idea of time must be a complex idea whose parts are in a temporal arrangement, and so must each of the ideas in its revival set. The exemplar of the idea of existence must be an idea of something taken to exist, as must each of the ideas in its revival set. In every case, however, these ideas will either be copied from an impression or composed of simpler ideas each of which has been copied from an impression, in accordance with the Copy Principle. Because each impression has multiple aspects of resemblance and can serve either as the exemplar or a member of the revival set of many different abstract ideas, a single simple or complex impression may be “of” many things: of Lassie, of collie, of dog, of brown, of space, of time, and of existence, for example, or of a flute, of a tune, of time, and of existence. There is no need for a separate impression “of” space or time that is distinct from the impressions of the things that appear in spatial or temporal arrangements, and no need for an impression “of” existence that is distinct from the impressions of the things that are perceived to exist.

### *The Separability Principle*

Hume’s second basic principle is sometimes called the *Separability Principle*: “Whatever objects are different are distinguishable, and . . . whatever objects are distinguishable are separable by the thought and imagination” (T 1.1.7.3). He uses this principle at several key points in the *Treatise* – for example, in arguing against the infinite divisibility of space, in arguing that there is no perception of time different from the perceptions of things in temporal succession, in arguing against the demonstrability of the principle that there is a cause for every beginning of existence, in arguing that there is no perceived “necessary connexion in nature” between causes and effects, and

in arguing that perceptions do not inhere in a mental substance. Although he refers to experience in remarking on how very readily the imagination can usually distinguish and separate, in idea, whatever things it finds to be different (T 1.1.3.4), the Separability Principle itself is evidently a direct consequence of his definition of the simple/complex distinction. For simple perceptions, as we have seen, are “such as admit of no distinction nor separation,” while the complex – that is, those that have parts that are different from one another – “are the contrary to these.” In offering this definition and the principle that is derived from it, Hume is of course assuming that things are separable if and only if they are distinguishable. This assumption is defensible, however, if the imagination can distinguish two things only by separating them in idea.

Despite the ultimate grounding of the Separability Principle in the definitions of “simple” and “complex,” it has often been objected (e.g., Kemp Smith 1941: 266; Mandelbaum 1974; Bricke 1980: 71) that Hume himself violates it in the discussion of *distinctions of reason* with which he concludes the section of the *Treatise* “Of abstract ideas” (T 1.1.7). He offers three examples of such distinctions: that between “figure [i.e., shape] and the body figur’d,” that between “motion and the body mov’d,” and that between “the figure and color” of a body. As he remarks,

the difficulty of explaining this distinction arises from the principle above explained, *that all ideas which are different are separable*. For it follows from thence, that if the figure be different from the body, their ideas must be separable as well as distinguishable; if they be not different, their ideas can neither be separable nor distinguishable. What then is meant by a distinction of reason, since it implies neither a difference nor separation? (1.1.7.17)

For Hume’s critics, the “difficulty” of reconciling these distinctions with the Separability Principle is an evident impossibility.

Hume, however, remarks that the difficulty can be resolved by appeal to the theory of abstract ideas that he has just explained:

It is certain that the mind would never have dreamed of distinguishing a figure from the body figured, as being in reality neither distinguishable, nor different, nor separable, did it not observe, that even in this simplicity there might be contained many different resemblances and relations. Thus, when a globe of white marble is presented, we receive only the impression of a white colour disposed in a certain form, nor are we able to separate and distinguish the colour from the form. But observing afterwards a globe of black marble and a cube of white, and comparing them with our former object, we find two separate resemblances, in what formerly seemed, and really is, perfectly inseparable. After a little more practice of this kind, we begin to distinguish the figure from the colour by a distinction of reason; that is, we consider the figure and colour together, since they are, in effect, the same and undistinguishable; but still view them in different aspects, according to the resemblances of which they are susceptible. (T 1.1.7.18)

The impression of the color of the entire globe of white marble is a complex impression composed of all of the spatially simple white impressions arrayed in certain shape. The impression of the figure of the globe is precisely the same complex impression composed of precisely the same components. This complex idea is neither different, nor distinguishable, nor separable from itself. The distinction is instead between two aspects in

which the white globe may resemble other objects; but these “aspects” are not *objects* in the sense of the Separability Principle (“whatever *objects* are different”). We recognize these aspects by means of two different abstract ideas: an abstract idea of white color, and an abstract idea of spherical shape. These abstract ideas *are* objects that are distinguishable and separable from one another – but they are also different, involving different revival sets, different token exemplars, and different terms. (For further discussion of the Copy Principle and the Separability Principle, see Garrett 1997: chs. 2 and 3.)

### *The Conceivability Principle*

A third principle of ideas, which may be called the Conceivability Principle, provides a criterion of possibility: “[N]othing we imagine is absolutely impossible” (T 1.2.2.8; see also T Abstract [Hume 1741], 11: “Whatever we conceive is possible, at least in a metaphysical sense”). Hume employs this principle in arguing that space and time can conform to our ideas of them, in arguing that any event may occur without a cause or without its usual effect, and in arguing that the annihilation of the soul is possible. He does not accept the unqualified converse of the principle – i.e., that what is inconceivable is impossible – since inconceivability may simply be the result of lacking the appropriate ideas, as will occur when one lacks the appropriate impressions. However, an inability to conceive, due to contradiction, something for which one has the appropriate ideas is a mark of impossibility.

While Hume characterizes the Principle of Conceivability as “an establish’d maxim in metaphysics,” it may also be derived from his account of necessity. For he holds that, in general, the application of the term “necessity” describes the mind’s determination to conceive things in a certain way – i.e., its inability to conceive and affirm otherwise – whether the necessity in question is that of “relations of ideas” (which are always either self-evident or demonstrable) or that of causes. Thus he writes:

As the necessity, which makes two times two equal to four, or three angles of a triangle equal to two right ones, lies only in the act of the understanding, by which we consider and compare these ideas; in like manner the necessity or power, which unites causes and effects, lies in the determination of the mind to pass from the one to the other.  
(T 1.3.14.23)

In the “absolute” necessity of relations of ideas, the inconceivability of the opposite is a function of the intrinsic character of the ideas themselves. Causal necessity is often conflated with absolute necessity in Hume’s view, but it is in fact a weaker kind of necessity, amounting not to the *absolute* unthinkability of the cause without its effect but rather in the psychological *difficulty* of separating their ideas following experience of their constant conjunction, together with a psychological *inability to believe* them to be separated. Given this understanding of necessity, the conceivability of something indeed entails its absolute possibility – subject only to the caveat that the mind must indeed conceive the very thing in question. To establish the possibility of squaring the circle, for example, it is not sufficient to imagine feeling pleased with oneself for having done so – one must imagine in detail how the procedure actually works.

### *The Principle of Association*

One of Hume's primary aims is to explain, as far as possible, why we have the complex ideas we do – both why simple ideas are combined into particular complex ideas at a single time and also why they compose the temporally complex trains of ideas that they do. In both cases, he invokes the *Principle of Association*: that ideas tend to occur in complexes when their objects are related by resemblance, contiguity, or cause-and-effect. (Employing the term “principle” in the somewhat different sense of “basic causal forces,” he also calls the three relations themselves “principles of association.”)

As one can infer from the parallels Hume offers in the Introduction to the *Treatise* concerning central figures in the development of the natural and the moral (i.e., human) sciences (T Introduction 7), it was his ambition to be the Newton of the moral sciences. The Principle of Association contributes to that aim particularly because, on analogy with gravitation, it provides

a kind of ATTRACTION, which in the mental world will be found to have as extraordinary effects as in the natural, and to shew itself in as many and as various forms. (T.1.1.4.6)

The Principle of Association explains why similar ideas are regularly combined into similar ideas of substances and modes by very different cultures, and why some particular ideas so often succeed other particular ideas in chains of thoughts, even the loosest reveries. Hume even invokes it to explain the requirements for “unity of action” in literary performances such as epic and dramatic poetry (EHU 1.3, “Of the Association of Ideas”). Most importantly, however, it proves to be of crucial importance in explaining probable reasoning, which is, on his view, by far the most frequent kind of reasoning. For all causal reasoning, he argues, depends on the relation of cause and effect. When ideas are related as ideas of causes and effects, the occurrence of one idea will immediately bring to mind the other; furthermore, if the first has the force and vivacity of a memory or belief, a share of that force and vivacity will be communicated to the idea of other, rendering it a belief as well. (This will also occur when the first perception is an impression.) Resemblance and contiguity may also contribute to the enlivening of one idea by another and hence contribute to the strength of a belief.

Hume was not the first to assign a role to the association of ideas in the explanation of human cognitive performance; Locke devotes a full chapter of his *Essay* (*Essay* II.33, “Of the Association of Ideas”) to the topic. But Locke appeals to the association of ideas exclusively to explain defective reasoning and madness: when ideas become associated together by “custom,” he holds, they become difficult to separate, and the tendency of one to draw the other with it leads to non sequiturs and fallacies. Proper reasoning for Locke, in contrast, depends on “perception of the agreement or disagreement” of ideas in the case of demonstrative reasoning; and in the case of probable reasoning, it depends on perception of the seeming or probable agreement or disagreement of ideas, mediated by the “grounds of probability” (which he identifies as “testimony” and “conformity to past experience”). Hume, in contrast, argues in effect that *all* probable reasoning is a kind of association of ideas produced by custom or habit. The perception of relations of agreement or disagreement among ideas is limited to the case of demonstrative reasoning.

It is because of this extensive use of the Principle of Association that Hume writes in his own anonymous review of the *Treatise*:

Thro' this whole book, there are great pretensions to new discoveries in philosophy; but if any thing can entitle the author to so glorious a name as that of an inventor, 'tis the use he makes of the principle of the association of ideas, which enters into most of his philosophy. (T Abstract 35)

In this respect, as in so many others, Hume's theory of ideas begins from a Lockean inspiration but transforms it radically so as to produce a foundation for his own philosophical "science of man."

See also 3 "Hume on Memory and Imagination"; 4 "Hume and the Origin of Our Ideas of Space and Time"

## References

- Bennett, Jonathan (1971). *Locke, Berkeley, Hume: Central Themes* (Oxford: Oxford University Press).
- Bricke, John (1980). *Hume's Philosophy of Mind* (Princeton, NJ: Princeton University Press).
- Flew, Antony (1961). *Hume's Philosophy of Belief* (London: Routledge and Kegan Paul).
- Garrett, Don (1997). *Cognition and Commitment in Hume's Philosophy* (New York: Oxford University Press).
- (2006). "Hume's Naturalistic Theory of Representation," *Synthese*, 152, pp. 301–19.
- Hume, David (1741). *An Abstract of a Book Lately Published: Entitled a Treatise of Human Nature* (London: C. Corbet). Included in Hume, David (1739–40), *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton (Oxford: Oxford University Press, 2000). (Cited by paragraph number.)
- Johnson, Oliver A. (1995). *The Mind of David Hume* (Urbana-Champaign, IL: University of Illinois Press).
- Kemp Smith, Norman (1941). *The Philosophy of David Hume* (London: Macmillan).
- Locke, John (1690). *An Essay concerning Human Understanding*, ed. Peter H. Nidditch (Oxford: Clarendon Press, 1979). (Cited by book, chapter.)
- Mandelbaum, Maurice (1974). "The Distinguishable and the Separable: A Note on Hume and Causation," *Journal of the History of Philosophy*, 12, pp. 242–7.
- Mounce, H. O. (1999). *Hume's Naturalism* (New York: Routledge).
- Owen, David and Cohon, Rachel (1997). "Representation, Reason, and Motivation," *Manuscripto*, 20, pp. 47–76.

## Further Reading

- Bennett, Jonathan (2001). *Learning from Six Philosophers*, vol. 2 (Oxford: Clarendon Press).
- Broughton, Janet (2006). "Impressions and Ideas," in Saul Traiger (ed.), *The Blackwell Guide to Hume's "Treatise"* (Malden, MA: Blackwell Publishing), pp. 43–58.
- Fodor, Jerry A. (2003). *Hume Variations* (Oxford: Clarendon Press).
- Fogelin, Robert (1984). "Hume and the Missing Shade of Blue," *Philosophy and Phenomenological Research*, 45, pp. 263–72.

- Noonan, Harold W. (1999). *Hume on Knowledge* (London: Routledge).
- Owen, David (1999). *Hume's Reason* (New York: Oxford University Press).
- Passmore, John A. (1952). *Hume's Intentions* (Cambridge: Cambridge University Press).
- Strawson, Galen (1989). *The Secret Connexion* (Oxford: Clarendon Press).
- Stroud, Barry (1997). *Hume* (London: Routledge and Kegan Paul).
- Waxman, Wayne (2005). *Kant and the Empiricists: Understanding Understanding* (New York: Oxford University Press).
- Wright, John P. (1983). *The Sceptical Realism of David Hume* (Manchester: Manchester University Press).