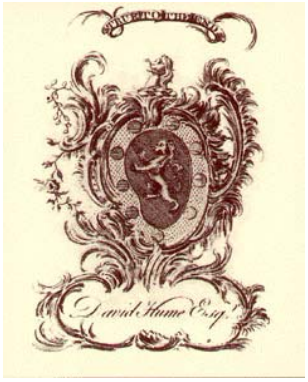

Filling the Gaps: Hume and Connectionism on the Continued Existence of Unperceived Objects

Mark Collier

Hume Studies Volume XXV, Number 1 and 2 (April/November, 1999) 155-170.

HUME STUDIES
Volume XXV, Numbers 1 and 2 , April/November 1999, pp. 155–170

Filling the Gaps: Hume and Connectionism on the Continued Existence of Unperceived Objects

MARK COLLIER

In Book I, part iv, section 2 of the *Treatise*, "Of scepticism with regard to the senses," Hume presents two different answers to the question of how we come to believe in the continued existence of unperceived objects.[1] He rejects his first answer shortly after its formulation, and the remainder of the section articulates an alternative account of the development of the belief. The account that Hume adopts, however, is susceptible to a number of insurmountable objections, which motivates a reassessment of his original proposal. This paper defends a version of Hume's initial explanation of the belief in continued existence and examines some of its philosophical implications.

The question of how we acquire the belief in continued existence poses a hard problem for Hume, since he is committed to two theses which severely constrain the answers he can give. The first thesis, indeed the first principle of Hume's science of human nature, is that all of our ideas are derived from impressions (T 7).[2] The second is that the sequences of impressions that constitute our acquaintance with objects are "gappy"; one need only turn one's gaze away from an object, or simply blink, to cause the train of perceptions to become fragmentary and interrupted.[3] The conjunction of these two theses threatens to render an empiricist explanation of continued existence intractable. On the one hand, the idea of continued existence must arise from the senses, yet on the other, the senses do not directly deliver this information.

Mark Collier is a Fellow in the Humanities at Stanford University, Department of Philosophy, Building 90, Stanford, CA 94305, USA.
e-mail: mcollier@stanford.edu

Because the senses can play no more than a partial explanatory role, Hume must recruit the help of an additional faculty to supplement the information they deliver. The faculty to which he turns, as he does so often in the *Treatise*, is the faculty of the imagination. Hume's "hypothesis" is that the belief in continued existence emerges from the interaction of the senses and the imagination, or in his own words, the "concurrence of some [sensory] qualities with the qualities of the imagination" (T 194). Hume's general explanatory strategy in *Treatise* I iv 2, then, is threefold. First, he will isolate the sensory qualities that are involved in attributions of continued existence. Second, he will provide an account of the principles of the imagination that accompany instances of the belief. Finally, he will explain how the principles of the imagination interact with the sensory qualities to produce the belief in continued existence.

In fact, Hume presents two formulations of this hypothesis. Let us reverse the order of Hume's exposition and begin with the second formulation. As dictated by his explanatory strategy, Hume must first choose the sensory qualities to play the partial explanatory role. Amid the changing contents of consciousness, he tells us, certain series of impressions exhibit the property of "constancy." A series of impressions is constant if it is interrupted, yet recommences without any qualitative alteration. Adopting Barry Stroud's formalism, we may express a paradigmatic constant series the following way, with letters representing impressions, and squares representing observational gaps:[4]

AAAAAA□□□AAAAAA

The resemblance of the impressions on each side of the gap is an important feature of constant series, since the resulting association leads the imagination to pass over the interruption in the sequence and conflate the broken series with one that is complete.[5] Subsequent reflection upon the appearance of such series, however, reveals their undeniable diversity. The result is a conflict of principles, which the imagination can only resolve through the supposition that the object continued to exist, although unperceived (T 199).

Unfortunately, there are a number of reasons why this explanation, which we might call the "conflation" account, has neither the virtues of plausibility nor consistency. First, the exclusive appeal to constancy represents a significant confusion on Hume's part. Constancy cannot be the sole sensory quality accompanying all of our ascriptions of continued existence, because we attribute continuity to both changing and unchanging series. Second, the appeal to a propensity of the imagination to disregard gaps between similar impressions appears intolerably ad hoc.[7] Finally, this formulation's employment of the imagination is in tension with Hume's portrayal of the automatic and implicit character of the imagination elsewhere in the *Treatise* (T 104). The conflation account requires the imagination to perform a metajudgment—the

resolution of conflict between judgments of unity and diversity—that should belong to the reflective faculty of the understanding.[8]

The deficiencies of the conflation account force us to reconsider Hume's original proposal, in the hope that it avoids these shortcomings. The first component of Hume's initial formulation is a sensory quality which he entitles the "coherence," or elsewhere simply the "regularity," of the changing contents of consciousness (T 195). Two sequences cohere when they conform to similar patterns of change. Unlike constancy, which is a characteristic of static objects like houses and chairs, coherence is a property of processes that fluctuate over time. Hume helps to clarify the meaning of the term coherence through his example of a gap in his observation of a fire. He remarks:

> When I return to my chamber after an hour's absence, I find not my fire in the same situation, in which I left it. (T 195)

Using the above formalism, we can describe Hume's impressions over this duration as:

AAABBB□□□DDDEEE

Hume notices that the qualities of the fire have changed from "B" to "D"; nevertheless, he continues:

> But then I am accustom'd in other instances to see a like alteration produc'd in a like time, whether I am present or absent, near or remote (T 195).

In other words, although the series is discontinuous, its pattern of change through time resembles a continuous series that has been previously observed:

AAABBBCCCDDDEEE

Hume's example is intended to bring out the fact that there are recurrent dependencies among the items of coherent series; for example, fires change from orange to gray, and not vice versa.

The important differences between Hume's two accounts, however, do not concern their choice of either constancy or coherence; once again, a comprehensive explanation of the belief in continued existence must cover both variable and invariable series. What really distinguishes the two formulations is their competing account of the role of the imagination in the development of the belief, and their rival proposals for the properties of the imagination that allow it to perform this function. Whereas the imagination resolves a contradiction between *identity* and *diversity* in the conflation account, in Hume's

original formulation it resolves conflicts between the *present* and the *past* (T 197). Once again, a look at one of Hume's examples will help clarify his reasoning. We are asked to imagine Hume still gazing at his fireplace, when suddenly he hears the sound "as of" a door turning upon its hinges. In the past, the sound of creaking door hinges had been accompanied by the sight of an opening door, but in the present instance the door is heard but not seen. Hume claims that this observation conflicts with the previous connection of the events, and in order to remove this disparity, the imagination fills in the gap by "supposing" that the door continues to exist unperceived (T 197). Hume maintains that this operation is completely commonplace:

> There is scarce a moment of my life, wherein there is not a similar instance presented to me, and I have not occasion to suppose the continu'd existence of objects, in order to connect their past and present appearances, and give them such an union with each other, as I have found by experience to be suitable to their particular natures and circumstances. Here then I am naturally led to regard the world as something real and durable, and as preserving its existence, even when it is no longer present to my perception. (T 197)

In order to complete this explanation, which we might call, with H. H. Price, the "assimilation" account, Hume needs only to articulate the principles of the imagination that underlie its capacity to fill in gaps.[8] Hume turns, then, to examine whether any of the resources of his science of human nature fit the bill. He begins with the habits of custom, but soon discovers that custom cannot play the required role, since custom cannot impose a greater regularity than is observed.

> '[T]is not only impossible, that any habit shou'd ever be acquir'd otherwise than by the regular succession of these perceptions, but also that any habit shou'd ever exceed that degree of regularity. Any degree, therefore, of regularity in our perceptions, can never be the foundation for us to infer a greater degree of regularity in some objects, which are not perceiv'd; since this supposes a contradiction, *viz.* a habit acquir'd by what was never present to the mind. (T 197)

Jonathan Bennett finds this dismissal puzzling, and suggests that Hume fails to provide adequate support for his claim that custom is insufficient.[9] Bennett, however, overlooks an important insight of Hume's regarding the limits of custom-based explanations of our belief in continued existence. Custom can explain why we anticipate seeing the door when we hear creaking hinges, since the two types of events have been constantly conjoined in the past, and we naturally infer from similar effects to similar causes (T 87). Nevertheless,

custom cannot explain why we believe the door continues to exist when this expectation is disappointed. As Hume makes clear in his analysis of probability, "contrary experiments" can weaken habits of custom, but habits of custom cannot strengthen contrary experiments (T 135). He concludes that the propensity to infer a greater regularity than is observed must be due to the influence of some other principles of the imagination (T 198).

With custom, the most likely candidate to serve as the desired principle of the imagination, deemed inadequate to perform the supplementative function, Hume turns to a principle of the imagination that he introduced in his discussion of the foundations of mathematics in *Treatise* I ii 4. In order to explain how mathematicians arrive at exact standards in geometry, Hume relied on the notion that the mind sometimes outruns the data of the senses and constructs imaginary standards, such as the standard of perfect quality (T 48). Hume now refers to this tendency of the imagination as a type of *cognitive momentum*.

> [T]he imagination, when set into any train of thinking, is apt to continue, even when its object fails it, and like a galley put in motion by the oars, carries on its course without any new impulse. (T 198)

Hume's galley metaphor, however, falls short of a principled account of the capacity of the imagination. His figurative language is more of an allusion to an explanation than an articulation of the precise qualities of the imagination that allow it to perform this role; worse, the appeal to a tendency to glide over interruptions appears at least as ad hoc as the principle invoked by the conflation account.

In any case, Hume is dissatisfied with his appeal to a bias of the mind toward the discovery of greater uniformity, and declares the principle "too weak to support so vast an edifice, as is that of the continued existence of all external bodies" (T 198–199). Hume also expresses dissatisfaction with his exclusive use of coherence, and claims that constancy must be added to the account.[10] So, without any further explanation, Hume drops his proposed solution, goes back to drawing board, and formulates his conflation account of the origins of the belief. Perhaps we ought not let Hume abandon his original proposal so easily, however. Bennett, for example, laments Hume's "abrupt" and "dogmatic" remarks about the weakness of the imagination, and maintains that Hume "turns his back on the success" of his initial explanation and "misrepresents it as a failure."[11] The interpretative question that ought to be pressed, then, is whether a Humean can rescue the assimilation account by demystifying the imagination's supplementative role in experience.

One proponent of an affirmative response is Price, who claims that Hume was on to something "real and important."[12] Although Price defends a version of Hume's account, he does so only after distinguishing between two ways in

which the imagination assimilates fragmentary observations. Price's first type of assimilation, "assimilation by convergence," captures the supplementary activity of the imagination in cases such as Hume's door example. In that example, a complete standard had been given to the senses in past experience. Price credits Hume with the insight that we do not jettison this standard, even in the face of recalcitrant experience; on the contrary, these observations are assimilated to our previous standards, and the order and structure of the past regularities are used to fill in their gaps.[13] Price refers to the capacity to impose past standards onto the present as the "inertia principle."[14] Unfortunately, Price does not attempt to identify the qualities of the imagination that account for this propensity, since he considers the inertia of the imagination to be an "ultimate and not further explicable tendency."[15]

Nevertheless, recent developments in cognitive science allow us to specify an associative mechanism that grounds the disposition to supplement incomplete perceptions. According to connectionist models of cognition, assimilation by convergence simply falls out of the way in which the mind processes information. Connectionist networks recognize objects by assimilating them to the prototype with which they share the most features.[16] Even when objects are partial or degraded, the network can still recognize them by a process known as "vector-completion," or what Paul Churchland calls "filling in the gaps."[16] Churchland provides an illustrative example of vector-completion by asking us to imagine a coyote who spots what appears to be a tail protruding from a dense patch of grass. Even though the rest of the animal is occluded, the coyote can complete the perceptual pattern after matching the tail to the prototype that it best fits.[18] Interestingly, Price uses a strikingly similar example in his account of the process of assimilation by convergence, when he asks us to imagine that we see a furry tail sticking out from behind a sofa.[19] We do not assume, he says, that the tail exists independently, but rather we fill in the perceptual gap by supposing that the tail is attached to an unobserved cat. In both vector-completion and assimilation by convergence, then, a complete standard is used to subsume partial sensory input, and this allows observational gaps to be filled in.

The hypothesis cannot rest on assimilation by convergence alone, however, since there may be cases where the past regularities used to assimilate broken series are themselves fragmentary. Although Hume himself does not consider this possibility, Price distinguishes a second type of assimilation appropriate for such cases, which he calls, following C. D. Broad, "assimilation by superposition." Assimilation by superposition involves the following procedure, which is general in the sense that it is indifferent to whether the series exhibits constancy (as in column 1) or coherence (as in column 2). First, one must observe a number of partial series, such as:

□AAAA	□BCDE
A□AAA	A□CDE
AA□AA	AB□DE
AAA□A	ABC□E
AAAA□	ABCD□

Although no continuous standards are observed, the imagination can produce a new, complete standard by superposing them one upon another:

AAAAA	ABCDE

Unlike assimilation by convergence, in which partial series are assimilated to past continuous series, in this type of case the past series are first assimilated to each other.[20] We may call the standards produced in this way "virtual standards." Since the virtual standards are complete, they can be used to fill in new instances of the partial sequences.

As was the case with assimilation by convergence, however, assimilation by superposition stands in need of defense. Which qualities of the imagination support the mutual assimilation of the past series? Without an account of the specific principles of the imagination that underwrite the alleged propensity, Price can be accused of unwarranted speculation. Once again, however, connectionist networks vindicate the assimilation account by grounding the disposition to superpose past standards in the intrinsic properties of an associative mechanism. In fact, connectionists themselves employ the term "superposition" to describe the way in which patterns of information are stored in networks.[21] Information storage is superpositional if the same resources (in this case, the configuration of weights across the units of the network) are used to represent all the patterns that are learned. Since multiple patterns are learned on the same set of weights, the network will develop a prototypical representation of their commonalities.[22] What this means, for our purposes, is that the network will automatically search for uniformity among the partial series that are observed and create virtual standards. Price's assimilation by superposition, according to connectionism, falls out of the way the mind stores information.

Once the assimilation account is extended along these lines, there remains nothing mysterious about the process whereby the imagination imposes a greater degree of regularity than is actually observed. Not only can the imagination fill in gaps in present sequences by assimilating them to past standards, but it can fill in gaps in the past standards themselves through the process of superposition. Moreover, this formulation of the hypothesis circumvents all three of the objections that were posed to the conflation account. First, both types of assimilation operate on either constancy or coherence; constant series are treated as a special case of coherent series in which all the items are mutu-

ally resembling.[23] Second, whereas the supplementative activity of the imagination appears ad hoc in the conflation explanation, in the assimilation account it is grounded in the fundamental properties of information-processing systems. Finally, unlike the conflation account's reliance on the imagination's ability to resolve contradictions, the process whereby the imagination fills in gaps in the assimilation account is nonreflective; all that is involved is the processing and storage of information in an associative mechanism.

Of course, it is one thing for a hypothesis to be consistently formulated, and another for it to be empirically adequate. Nonetheless, not only does connectionist theory allow us to complete the assimilation hypothesis by supplying the missing principles of the imagination, but connectionist methodology provides the experimental conditions under which the hypothesis can be implemented and tested. Connectionist researchers in the field of cognitive development have recently devised computer models that simulate the acquisition of a child's concept of object permanence.[24] The researchers use a Simple Recurrent Network (SRN) for their experiments (see Figure 1).[25]

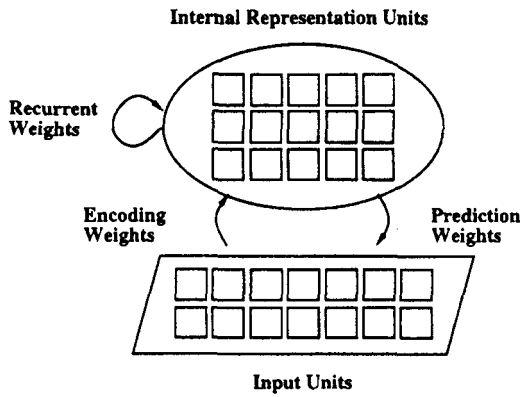

Figure 1: Simple recurrent network architecture

SRNs are essentially inductive mechanisms; they learn to make predictions about what will happen in the future on the basis of what has happened in the past. SRNs also perform vector-completion and superposition. These networks thus serve nicely as models of the mental machinery posited by the assimilation hypothesis.

In the connectionist simulations, the researchers present the SRN with a stream of images projected onto an artificial retina. These sequences represent what developmental psychologists refer to as an "occlusion event", in which an object disappears and reappears from behind an occluder.[26] An illustration of the type of sequences used in these experiments is provided in Figure 2.[27]

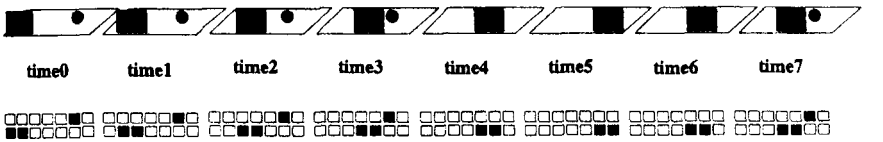

Figure 2: Simulation training data

We can interpret the event sequence in Figure 2 in terms of our formalism by letting "A" stand for a time step where the object is visible, and letting a gap correspond to a time step during which the object is occluded:

AAAA□□□A

Notice that this a constant series, since the object does not change after the occlusion. When the network is presented with such sequences, the problem it confronts is a formal analogue of the one that Hume faces when he "shuts his eye" or "turns his head" and loses sight of the invariable "mountains, and houses, and trees, which lie at present under my eye" (T 194). According to the assimilation account, though, what does the explanatory work is not the mutual resemblance of the items in the sequence, but the resemblance between the sequence and complete standards observed in the past.

The assimilation hypothesis predicts that the interaction of the sensory data with the principles of assimilation will be sufficient to infer continued existence. Let us now turn to the simulation results and see whether they substantiate this prediction. How well do these networks learn the task? During the early stages of training, the network shows little success. The reason is that the network bases its inferences on the past behavior of the object, and at this point in development it has had very little experience. In other words, the network has no a priori understanding of the characteristics of objects, and without having sampled a large enough portion of the data set, the network would have no way to anticipate that a particular train of events will occur. It is only by observing the contingencies of the data that the network can learn to make successful predictions.

Once the network has had sufficient exposure to the regularities in the data set, it does learn to anticipate that the object will be visible when the occluder moves away. The learning curve in Figure 3 demonstrates that the accuracy of the network's inferences about the continued existence of objects develops gradually as a function of its increasing experience.[28] When the training period is complete, the network predicts the object will be visible when the barrier moves away with a small degree of error. According to the connection-

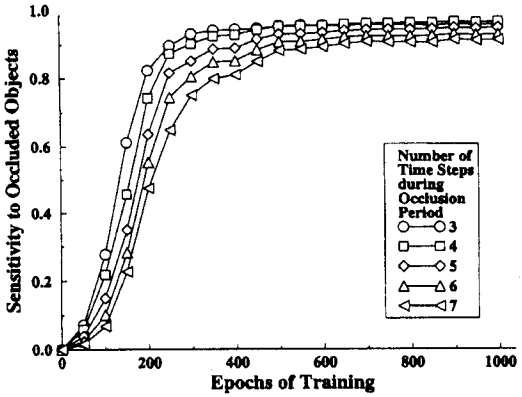

Figure 3: Simulation learning curve

ist researchers, the trained network has developed a "sensitivity" to the continued existence of the object.[29]

Of course, one might argue that sensitivity to the reappearance of the object is not itself a guarantee that the network represents the object as continuing to exist while unperceived, since the network might solve the prediction task by inferring that the object pops in and out of existence during the occlusion event. How can we be sure the network has in fact learned to represent the *continued* existence of the unperceived object in order to solve the task? One of the relative benefits of computer simulation over human experimentation is that modeling offers researchers a chance to peek into the head, as it were, of the network in order to examine the representations that are formed during learning.[30] In their simulation, the researchers recorded the patterns of activity across the network's hidden unit representations of the object as it learned to predict its re-emergence from behind the occluder.[31] Here are the Hinton diagrams of these units (1, 8, 10, 11, and 15) after 100, 200, and 1,000 epochs of training. The units that code for the object are shown as shaded squares, and the darkness and size of the square correspond to the magnitude and sign of the connection.[32] Notice the gradual increase of the strength of the internal representation of the object during those time steps when it is occluded.[33]

As the assimilation hypothesis predicts, the network can extract continuity from a discontinuous data set, and it can do so without relying on any innate representations. The experimental results demonstrate that the belief in continued existence can arise solely from the interaction of sensory information with the principles of an information-processing mechanism. The simulations are therefore "existence proofs" of the possibility of an empiricist solution to the problem with which we began.

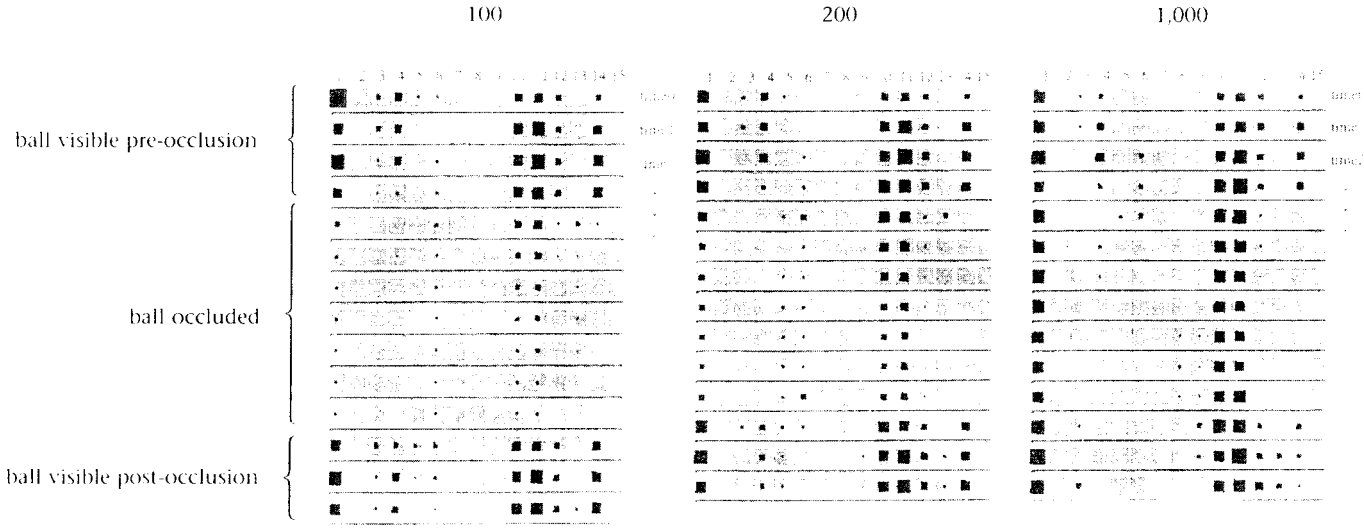

Figure 4: Hinton diagram of the network's hidden units

The idea that connectionism vindicates Hume might appear puzzling at first. How, after more than 250 years, could a research program in cognitive science converge on a similar answer as Hume? The surprise dissipates, however, upon recognition that Hume and connectionism confront the same problem space—to give a causal explanation of the origins of our belief in continued existence—and they impose similar constraints on its solution. First, connectionists agree with Hume that our acquaintance with objects is often punctuated by interruptions. Second, connectionists hold a version of Hume's thesis that the representation of continued existence must nonetheless be derived from the sensory input. While connectionism and nativism are not logically incompatible—a connectionist could easily handset the initial weights of the network—their account of development is offered as an *alternative* to the view that innate representations are required to guide the acquisition of knowledge of continued existence.[34] Since they make these two Humean commitments, connectionists face the Humean challenge of explaining how continued existence can be inferred from discontinuous data, which makes an appeal to principles that allow the network to transcend the data a sensible route.

Given the recent evidence from cognitive science, we might agree with Bennett's assessment of Hume's initial explanation: "Schematic as the account admittedly is, it seems to be on the right lines."[35] With hindsight, it is unfortunate that Hume prevented himself from thinking through his original proposal. Hume abandoned his assimilation hypothesis because he lacked a viable principle of the imagination that could account for its capacity to fill in gaps. Nonetheless, recent evidence from cognitive science vindicates Hume's appeal to a supplementative propensity of the imagination, and thus provides his account of the etiology of the belief in the continued existence of unperceived objects with a great deal of plausibility.

Let us end our examination of the assimilation account by considering some of its philosophical implications. Hume's purpose in presenting his psychological explanation of continued existence is to account for our ordinary, or in Hume's terms vulgar, beliefs about objects. According to the conflation explanation that Hume adopts, it turns out that these beliefs harbor a number of confusions. As a result, Hume closes the section "Of scepticism with regard to the senses" by drawing a number of "despairing conclusions" about the belief in continued existence.[36] These conclusions are premised on his conflation account, however, and the epistemological status of the natural belief in continued existence must be revised when we adopt the assimilation explanation.

Hume's conflation account maintains that the belief in continued existence results from a tendency of the imagination to "confound" gappy and complete series (T 203). If we adopt the assimilation account, though, the imagination is no longer predisposed toward such an error, since the assimilation of fragmentary sequences to complete past standards is not a case of mis-

taken identity, but one of best fit. Moreover, according to the conflation account, the imagination resolves the conflict between the judgments of identity and diversity by supposing that impressions exist unperceived. On the assimilation account, however, the vulgar are involved in no such muddle. According to that account, impressions cease to exist the moment we lose consciousness of them, but traces of these impressions remain in memory and can be used to fill in gappy sequences.

Another consequence of adopting the assimilation account is that the wind is taken out of the sails of the "philosophical system."[37] The reason is that the philosophical system is parasitic on the confusions of vulgar consciousness; the philosopher goes to great lengths to show that, contra the vulgar, sensory impressions are in fact "interrupted and perishing" (T 210–211). When we embrace the assimilation account, then, the critique of the philosopher misses its target. Moreover, when we abandon the conflation account, the metaphysics of the philosophical system is no longer motivated. The philosophical system attempts to reconcile the gappiness thesis and the belief in continued existence by inventing the "double existence" of mind-dependent perceptions and mind-independent objects, and "ascribing the *interruption* to perceptions, and the *continuance* to objects" (T 215). According to the assimilation hypothesis, however, the vulgar need not venture beyond the "universe of the imagination" to form beliefs about the continued existence of objects, since gaps in fragmentary series can be filled by items similar in kind to those with which we have been acquainted (T 68).[38]

Hume tells us that he began *Treatise* I iv 2 with trust in his senses, but as a result of "reviewing" his conflation account of the belief in continued existence, he is now "of a quite contrary sentiment, and more inclin'd to repose no faith at all in my senses, or rather imagination, than to place in it such an implicit confidence" (T 217). This consequent skepticism comes to a culmination on the last page of the section, where Hume writes:

> What then can we look for from this confusion of groundless and extraordinary opinions but error and falsehood? And how can we justify to ourselves any belief we repose in them? (T 218)

These confusions, however, are artifacts of the conflation explanation, and when we view the natural belief in continued existence through the lenses of the assimilation account, it no longer results from a propensity towards error, and no longer gives rise to any groundless convictions about the existence of unperceivable objects. Skepticism with regard to the senses must therefore find some other foothold.

NOTES

Figures 1–4 in this article are reproduced from Y. Munakata, J. McLelland, et al., "Rethinking Infant Knowledge: Toward an Adaptive Process Account of Successes and Failures in Object Permanence Tasks," *Psychological Review*104 (1997), 699, 700, and 702–3. I am grateful to Yuko Munakata and the editors of *Psychological Review* for permission to reproduce them here.

1. Of course, Hume is concerned in this chapter with the origins of our belief in the continued and *distinct* existence of bodies, where distinctness is itself a complex term referring to the external and independent existence of bodies. We shall limit ourselves here to a discussion of continued existence, however, since the question of continuation is logically prior to the question of distinctness. Although Hume suggests that continuity and distinctness are materially equivalent, a little reflection reveals that although the question of continuation decides the question of distinctness, the converse does not hold, since there are cases of distinct yet discontinuous processes, such as a series of lightning flashes in a thunderstorm. (See H. H. Price, *Hume's Theory of the External World* [Oxford: Oxford University Press, 1963], 18.)

2. References are to D. Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, 2nd ed. revised by P. H. Nidditch (Oxford: Clarendon Press, 1978), abbreviated by "T" followed by the page number in that edition.

3. Price, *Hume's Theory of the External World*, 20.

4. Barry Stroud, *Hume* (London: Routledge & Kegan Paul, 1977), 101.

5. Hume attempts to clarify his account in a later footnote. He tells us that there are in fact two resemblances at work: first, there is the resemblance between the items of constant series; this causes the imagination to pass smoothly from one element of the series to the next; second, this smooth passage of the mind resembles, and is thus confounded with, the passage of the mind as it "surveys" complete series (T 205 n.).

6. The worry is that the law of resemblance explains why the various A's are associated, but not why they are conflated. The same worry applies to the resemblance between the mind's passage over complete and gappy series; the resemblance explains why we associate these two types of series, but not why we confound them. The tendency to conflate resembling series is an additional propensity which Hume employs solely for the purpose of explaining the origins of our idea of identity.

7. Stroud, *Hume*, 108–109. The worry is that Hume commits a category mistake when he speaks of an associative faculty as "resolving conflicts". Hume's appeal to a process of conscious reflection stands in tension with his Hutchesonian project of treating belief as a species of sensation. Moreover, since Hume's theory of the imagination is supposed to have a general application to animals and infants, he can only appeal to a process of reflection at the cost of greatly narrowing the scope of his theory (T I iii 16).

8. Price, *Hume's Theory of the External World*, 72.

9. Jonathan Bennett, *Locke, Berkeley, Hume: Central Themes* (Oxford: Oxford University Press, 1971), 328.

10. This realization is one of Hume's rare careful moments on this point, and when he turns to his second formulation of his hypothesis, coherence drops out of the picture completely.

11. Bennett, *Locke, Berkeley, Hume*, 329–330.

12. Price, *Hume's Theory of the External World*, 59.

13. Ibid., 58.

14. Ibid., 54.

15. Ibid., 58.

16. See Andy Clark, *Associative Engines: Connectionism, Concepts, and Representational Change* (Cambridge, Mass.: MIT Press, 1993), ch. 2 for a clear overview of prototype recognition in connectionist networks.

17. Paul Churchland, *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain* (Cambridge, Mass.: MIT Press, 1995), 280. *Technical Note: Vector Completion.* Vector completion is an instance of what is known in cognitive science as a "familiarity effect." The reason why networks are capable of filling in missing information is that connectionist representations are distributed over a set of interconnected units. Thus, when part of the pattern is presented as input, the units that are turned on will excite the units which code the rest of the pattern.

18. Paul Churchland, *The Neurocomputational Perspective* (Cambridge, Mass.: MIT Press, 1989), 211.

19. Price, *Hume's Theory of the External World*, 71.

20. Ibid., 75.

21. Clark, *Associative Engines*, 17. Clark considers superposition part of the "USP," or Unique Selling Point, of connectionism.

22. Ibid., 20.

23. Price, *Hume's Theory of the External World*, 60. Price argues that these qualities are two aspects of a more general principle that he calls "Gap Indifference." See also Bennett, *Locke, Berkeley, Hume*, 323.

24. Y. Munakata, J. McClelland, et al., "Rethinking Infant Knowledge: Toward an Adaptive Process Account of Successes and Failures in Object Performance Tasks," *Psychological Review* 104 (1997): 686–713.

25. Munakata, McClelland, et al., "Rethinking Infant Knowledge," 699. *Technical note: Simple Recurrent Networks.* The network is recurrent because information not only flows from the sensory input layer to the hidden layer (where internal representations are formed), but information also flows back down from the hidden layer to the input layer. This recurrent connection provides the network with short-term memory and allows it to learn sequences that unfold over time. These networks are dynamical systems, because the state of the network at time t is a function of its state at time t–1. Like most connectionist networks, SRNs learn by error correction and back propagation. That is, if the prediction of the next time

step is inaccurate, the difference between the prediction and the target pattern is computed, and this value is used to change the weights (in which the networks' knowledge is stored) in order to allow better predictions in the future. An important feature of SRNs is that they do not need an external "teacher" to determine the target pattern, since this pattern is simply the next sensory input.

26. Jeff Elman, Elizabeth Bates, et al., *Rethinking Innateness: A Connectionist Perspective on Development* (Cambridge, Mass.: MIT Press, 1996), 51.

27. Munakata, McClelland, et al., "Rethinking Infant Knowledge," 700.

28. Ibid.

29. Ibid, 701.

30. Elman, Bates, et al., *Rethinking Innateness*, 45; Kim Plunkett and Jeff Elman, *Simulating Nature, Nurture: A Handbook of Connectionist Exercises* (Cambridge, Mass.: MIT Press, 1997), 20.

31. *Technical Note: Methods of Analysis.* In order to distinguish the network's representation of the hidden object, the researchers used the method of "stimulus subtraction." "[T]o isolate the network representation of the ball during events involving a barrier, we record the pattern of activity across the network's internal representation units at a particular time step in a particular 'ball-barrier' event and subtract from it the pattern of activity from the corresponding time step in the corresponding 'barrier-only' event" (Munakata, McClelland, et al., "Rethinking Infant Knowledge," 700).

32. Plunkett and Elman, *Simulating Nature, Nurture*, 30.

33. Partial reproduction of Munakata, McClelland, et al. "Rethinking Infant Knowledge," 702–703.

34. Elman, Bates, et al., *Rethinking Innateness*, 150.

35. Bennett, *Locke, Berkeley, Hume*, 1971, 327.

36. Ibid.

37. Hume clearly has Locke's Representational Realism in mind when he refers to the "philosophical system."

38. See Price, Hume's Theory of the External World, 93: "The supplementations which we postulate are just *continuations* of our sense-impressions, homogeneous with the data whose continuations they are taken to be." Indeed, one result of the connectionist simulations was that the network's representations of visible and occluded objects became more similar with the increase of experience: Munakata, McClelland, et al., "Rethinking Infant Knowledge," 704.