Philosophy of Science (2024), **91**, 1296–1305
doi:10.1017/psa.2023.139

PSA PHILOSOPHY OF SCIENCE ASSOCIATION

PHILOSOPHY OF SCIENCE

SYMPOSIA PAPER

Hume's Externalist Gambit

Hayley Clatterbuck

Department of Philosophy, University of Wisconsin-Madison, USA
Email: clatterbuck@wisc.edu

(Received 20 April 2023; revised 29 September 2023; accepted 05 October 2023; first published online 20 October 2023)

Abstract

I examine three arguments that purport to show that connectionist, associationist architectures cannot achieve key features of human thought. Hume anticipated each of these three arguments and provided a unified strategy for responding to each, the "externalist gambit." On this account, external natural language provides the necessary structure for associationist systems to achieve those features of thought that their opponents take them to lack. The externalist gambit provides a promising avenue for today's defenders of connectionism about the human mind.

1. Introduction

The surprisingly successful performance of Deep Neural Nets (DNNs) across many cognitive domains—including game playing, image recognition, and text generation—has reignited old debates about the architecture of the mind, as these systems resemble, in key aspects, associationist architectures proposed by British empiricists and more recent connectionists. The analogy between empiricist theories of the architecture of the mind and modern-day DNNs can bear philosophical fruit in both directions. First, if we see DNNs as implementations of the associationist picture—albeit using technological and statistical tools undreamt of by the British empiricists—then they may constitute an empirical test of the theory. Second, those who defend DNNs as a model for human cognition can look to concepts and arguments developed by earlier associationists to defend against their philosophical detractors.

Here, I will briefly defend an analogy between Hume's associationist psychology and DNNs. I will then examine three more recent arguments that purport to show that connectionist, associationist architectures (including DNNs) cannot achieve key aspects of human thought. I will then show that Hume anticipates each of these arguments. In responding to each, Hume utilizes a strategy that I will call the *externalist gambit*, according to which natural language (not an internal language of thought) provides the structure that associationist architectures are purported to lack. Lastly, I will argue that Hume's externalist gambit provides a single, unified strategy for defending DNN architectures as models of the human mind.

2. The analogy

The debate between associationists and their critics has persisted for centuries. Major flashpoints include British empiricism in the eighteenth century, the connectionist movement of the 1980s and 1990s, and the advent of extraordinarily successful deep-learning algorithms in the past few years. While there are certainly differences among these eras and among researchers within each movement, they are united by similar basic commitments about the structure of the mind. Indeed, some of the most preeminent developers of modern DNNs have explicitly pitched their work as being in the tradition of the British empiricists (Silver et al. 2017).

At the crux of the associationist picture is the claim that (probabilistic) causal connections among (perceptual) representations constitute the basic architecture of the mind:[1]

> Associationism is a theory that connects learning to thought based on principles of the organism's causal history. Since its early roots, associationists have sought to use the history of an organism's experience as the main sculptor of cognitive architecture. In its most basic form, associationism has claimed that pairs of thoughts become associated based on the organism's past experience. (Mandelbaum 2020)

For our purposes, the most important commitment of associationism is that there is little structure to the mind apart from the structure that comes from experience.[2]

Contrast this position with the Language of Thought (LoT) hypothesis, which holds that thought has structure in the same way that natural language has structure. Sentences in a natural language have grammatical structure; they are not mere concatenations of associations. Their semantic parts are discrete and have stable meanings across contexts, and syntax provides rules for combination into new complex wholes. For example, "Mary loves John" does not merely assert an association between Mary, loving, and John, for that association is equally compatible with "John loves Mary." For LoT theorists, the causal transitions between thoughts do not merely reflect their cooccurrence in experience but they also reflect the rational and semantic relations between the contents of those thoughts. Importantly, this structure is not (and cannot be) derived from experience.

In this article, I will be focusing on Hume's associationist architecture and objections to associationism from LoT theorists. My thesis will be that Hume anticipated these objections and that his strategy for responding to can be of service to modern-day associationists. While a full treatment of Hume's view is not possible here, and treatments of connectionism and modern machine learning are even less forthcoming,[3] a brief statement of the analogy will have to suffice.

[1] I say "basic" because associationists have sometimes posited various faculties in addition to association. For example, Hume posited the faculty of imagination, which according to Fodor and Pylyshyn (1988 n. 29), departs significantly from the core of associationism. This faculty approach is explored in depth by Buckner (2023).

[2] Of course, no theory can completely eschew innate structure. Even for staunch empiricist associationists, the principles of association must be part of the structure of the mind.

[3] For a helpful introduction, see Buckner (2019).

https://doi.org/10.1017/psa.2023.139 Published online by Cambridge University Press

According to Hume, all ideas are copies of sensory impressions. Two ideas, A and B, will become associated if (and exempting the imagination, only if) their corresponding impressions were experienced contiguously, as being similar, or standing in cause-effect relationships. The strength of the association—the strength of the disposition for the idea of A to elicit the idea of B—will be determined by the frequency with which A and B were experienced together in one's past or their similarity.

DNNs consist of a set of linked nodes. The weight of a link between reflects the association between two nodes, and these weights change with experience. An input layer of nodes registers features of experience (say, the color values of each pixel in a photo of a dog), an output layer generates some behavior (say, a prediction that the photo is of a dog), and interior layers process associations among features of the input and eventual output. Through reinforcement learning, weights are adjusted until the output behaviors match the learning target. The statistical tools used to change these weights are much more sophisticated and diverse than those envisioned by Hume, but the architecture shares the same general associationist commitments: Structure consists of associations among representations and those associations come from experience.

3. The externalist gambit

A key commitment of Hume's associationism is the thesis that relations among thoughts are caused by corresponding relations among experiences. It is in this sense that Hume's theory is *externalist*. The associating force between two ideas is not to be found in some *internal* or *intrinsic* property of the ideas, but rather in the *external* fact that the one idea tends to cause the other, in virtue of how they were related in experience.

If external relations are sufficient to explain why ideas become associated, we need not posit internal representations of the relations they bear to one another. Consider the fact that an idea of lightning tends to elicit an idea of thunder. An internalist explanation will posit that we represent lightning, thunder, *and* the relation of contiguity they bear to one another. The externalist argues that we represent lightning and thunder, and contiguity is a fact about how the ideas are related in the mind, not a third idea. Hume is most famously and forcefully an externalist about the relation of causation,[4] arguing that there is no internal representation of a necessary connection between cause and effect; instead, all that exists is the external fact that ideas of the cause tend to bring about ideas of the effect. More generally, Hume's externalist strategy is to explain certain properties of thought (e.g., that we expect effects to follow their causes) without positing that there is some internal representation of that property (e.g., an idea of a necessary connection between cause and effect).

My focus here is how he uses this strategy to argue against the claim that there are some features of thought that the mind can possess only if it has an *intrinsic linguistic* structure, whether this takes the form of representations of grammatical rules or a nonrepresented, language-like functional architecture (Pylyshyn 1991). As an externalist, Hume must explain the structure of thought by locating a corresponding

[4] Elsewhere, Hume seems to posit that we sometimes do have ideas of relations that are derived from experience (Inukai 2010; Clatterbuck 2016).

https://doi.org/10.1017/psa.2023.139 Published online by Cambridge University Press

structure in experience. He does so by arguing that experience with natural language provides a set of regularities in the world that our minds pick up on and that provides linguistic structure to our thought. Thus, to the extent that thought has the properties of language, this comes from outside in rather than inside out.

This externalist gambit has become an increasingly popular strategy among philosophers and cognitive scientists who deny that the mind has intrinsic linguistic structure. For example, Clark (2006, 293) argues that natural language is an external technology that helps augments our more connectionist cognitive architecture "without installing any fundamentally new styles of representation or processing within that machinery." Machery (2005) argues that inner speech is a sensory reenactment of auditory stimuli. Our thought has linguistic-like structure only extrinsically, in virtue of being about linguistic objects (natural language sentences), but "this does not license any inference that the thought itself possesses" those linguistic properties (477). A clear statement of the view comes from Lupyan and Bergan (2016, 414):

> There is no need for a language of thought. It is not that we think "in" language. Rather, language directly interfaces with the mental representations that are used in perception and action, helping to form the (approximately) compositional, abstract representations that thinkers like Fodor take as *a priori*.

4. Three objections to associationist architectures

In what follows, I will consider three objections that LoT theorists have made against connectionist architectures; three features (F) they allege that human thought has but that associationist architectures do not. Hume anticipates each of these objections and, in his response to each, he:

a. Denies that F is a genuine or universal feature of human thought;
b. Argues that there is more F in associationist systems than his opponent claims; and
c. Argues that external natural language provides the excess structure to provide thought with F.

4.1. Abstractness

The first objection is that Hume's proposed architecture cannot achieve abstract or general thought. A central premise of Hume's system is that all ideas are copies of sensory impressions (or combinations thereof). However, we have some ideas that do not seem to reduce to sensory impressions. The problem of abstraction can be extrapolated to associationism more generally. A truly abstract, general idea applies to an entire category of things, for example, a thought about *triangularity*. However, if thoughts are mere associations of experience, then an idea of triangles will be, at best, a statistical summary of triangles that one has already seen. Therefore, it will fall short of true generality.

https://doi.org/10.1017/psa.2023.139 Published online by Cambridge University Press

Hume spends a considerable amount of Book I of the *Treatise* responding to this objection, and his response to it will form the basis of his replies to further objections. There are, in fact, three related problems of abstraction. First, some properties are never experienced on their own. For example, we never encounter an object's color independently of its shape. Thus, it's unclear how we can form an idea of REDNESS that is abstracted from the variously shaped objects that are red. Second, experienced objects have determinate properties, but we possess ideas that abstract away from such details. For example, we have an abstract idea of TRIANGLE that is neither isosceles nor acute and that has no particular length or color. However, every triangle we experience has determinate properties (e.g., is isosceles, 2 inches long, and blue). Thus, the abstract idea cannot be copied from any sensory particular, nor can it be a combination of them. Third, some ideas (e.g., DEMOCRACY) do not seem to be definable in terms of any perceptual properties or regularities at all.

How, then, do operations over perceptual inputs give rise to representations of such abstract notions? Hume's first answer is that they do not, at least not in the *internalist* sense. Just as we do not have an idea of the necessary connection between cause and effect, we do not have an idea of that abstract property in virtue of which, for instance, all triangles are of a kind. Instead, when we have a general thought about triangle-hood, what we have is an idea of a particular triangle (e.g., a blue isosceles) and a disposition to bring to mind a series of ideas of other particular triangles (e.g., a green equilateral, a black right triangle).

The abstractness or generality of an idea, then, is not an intrinsic property of it but rather an external relation that it bears to other ideas (here, the disposition to elicit other ideas of objects of the same class). Hume takes his inspiration here from Berkeley (1975, 6), who argues, "the only kind of universality that I can grasp doesn't belong to anything's intrinsic nature; a thing's universality consists how it relates to the particulars that it signifies or represents." As Hume (1978, 20) puts it:

> Abstract ideas are therefore in themselves individual, however they may become general in their representation. The image in the mind is only that of a particular object, tho' the application of it in our reasoning be the same, as if it were universal.

To have an abstract thought about triangles, then, is to have a particular kind of disposition, a disposition for the idea of a particular triangle to conjure ideas of other triangles. How does a particular idea come to have this disposition?

Hume's answer is that natural language provides the necessary structure for a particular idea to become general:[5]

> A particular idea becomes general by being annexed to a general term; that is, to a term, which from a customary conjunction has a relation to many other particular ideas, and readily recalls them in the imagination . . . . The word raises up an individual idea, along with a certain custom, and that custom produces any other individual one for we may have occasion. (ibid., 22)

[5] For an alternative, Lockean account of how DNNs can achieve abstraction, see Buckner (2018).

https://doi.org/10.1017/psa.2023.139 Published online by Cambridge University Press

Our idea of the word "triangle" is yet another idea of a particular sense impression (the sound or sight of the word). Yet this word has become associated with a diverse array of triangles, providing the glue that unites them into a class. Hence, when a particular triangle is thought of in conjunction with its label, that label elicits the associations with all of the other triangles that have been paired with the word.

This raises a worry for Hume's account. Suppose "triangle" only causes us to think of those particular triangles have been paired with the label in our experience. Because this is a limited class, our thoughts about triangles will not achieve true generality. It would not permit us to extend our knowledge or reasoning about triangles beyond the exemplars that we have seen.

In response, Hume argues that the pairing of a class of objects with a label can change how that class of objects is represented, making it more general. First, the class of items associated with a word (red things with "red," triangles with "triangle") will typically share some aspect of resemblance. Their association with a common label causes this aspect of resemblance to become more salient, such that the word will elicit ideas of objects that also share this feature (Hume 1978, 25):[6]

> As the individuals are collected together, and placed under a general term with a view to that resemblance, which they bear to each other, this relation must facilitate their entrance in the imagination and make them be suggested more readily upon occasion. (ibid., 23)

Second, as we learn more and more associations between general terms and classes of items, we learn a higher-order association: Labels are associated with general classes. Now, when even a single item is associated with a general term, we will seek to use that single item in a general rather than particular way (ibid., 105).

In this way, Hume argues that natural language provides the necessary structure for abstract thought. Labels are the string that tie individuals together and make them cohere into a general kind. Hume's insights have been explored in experimental work by researchers inclined toward the externalist gambit (Smith and Heise 1992). For example, Lupyan et al. (2007, 1082) showed that subjects were more successful in making learning abstract classes when given labels, because "as a label is paired with individual exemplars, it becomes associated with features most reliably associated with the category. When activated, it then dynamically creates a more robust category attractor." Likewise, it has been argued that children learn the higher-order association that labels indicate general kinds, that labels are "invitations to form categories" (Waxman and Markow 1995, 257).

4.2. Systematicity of meaning

A second objection to associationist theories is that they cannot explain the systematicity of human thought. There are two aspects of systematicity: meaning and

[6] One might object that if we associate two ideas in virtue of an apprehension of their similarity, it appears that internal properties of the representations are doing the work. Bradshaw (1988, 18) argues that this association by similarity is compatible with Hume's externalism, for "one does not have an idea of resemblance ... similarity or resemblance is taken to be a fact which contributes to an association, not an object of any idea."

https://doi.org/10.1017/psa.2023.139 Published online by Cambridge University Press

inference. Systematicity of meaning is the claim that there are systematic relationships between thoughts that a thinker can entertain that result from similarities in their meaning. For example, anyone who can think (and understand) "John loves Mary" can also think "Mary loves John." Because it posits recurring discrete parts that are combined in accordance with systematic rules, the LoT can explain systematicity. Anyone who can think the first thought has the parts "Mary," "John," and "loves" and a rule for combining them into a meaningful whole. These ingredients suffice to generate the second thought.

However, for Hume, every thought about say, triangles, is a thought about a particular triangle. Therefore, there is no guarantee that you think the same thought when you think about triangles in different contexts. Given that one's thoughts about triangles are determined by one's experience, it seems possible that a person could think thoughts about blue isosceles triangles but not red right ones. One's knowledge of triangles may also be distributed; you might have learned that the angles add up to 180° on the basis of isosceles triangles and how to calculate their area from right triangles. This generates the following argument against Hume's system:

1. "Sentences that are systematically related are composed from the same syntactic constituents . . . . Insofar as a language is systematic, a lexical item must make approximately the same semantic contribution to each expression in which it occurs." (Fodor and Pylyshyn 1988, 38)
2. Representations in associationist architectures are context sensitive and distributed.
3. If (2), then lexical items in associationist architectures will not make approximately the same semantic contribution to each expression in which it occurs.

C: Therefore, associationist architectures will not be systematic.

What makes all our thoughts about triangles cohere into a unified concept? How can we make sure that all our ideas of triangles bear enough similarity to be reused across contexts? Hume's first reply is to note that there will often be real resemblances among associated items. Hence, even if (2) is true (3) does not follow. If every idea of a triangle has certain features in common, then each triangle thought can make a similar semantic contribution.[7]

His second response is to reutilize the externalist gambit. When we think about a particular triangle, we draw inferences from that particular. However, when it is paired with a general term, it becomes general in its signification. It will elicit a train of ideas of other triangles that have also been associated with the word, and we can use this distributed set to find those properties that are common to triangles in general:

> After the mind has produced an individual idea, upon which we reason, the attendant custom, revived by the general or abstract term, readily suggests any other individual, if by chance we form any reasoning that agrees not with it.

[7] This is akin to Shea's (2007) argument that clusters of activation in connectionist networks, rather than nodes, are stable representations.

https://doi.org/10.1017/psa.2023.139 Published online by Cambridge University Press

> Thus should we mention the word, triangle, and form the idea of a particular equilateral one to correspond to it, and should we afterwards assert that the three angles of a triangle are equal to one another, the other individuals of a scalenum and isosceles, which we overlooked at first, immediately crowd in upon us, and make us perceive the falsehood of this proposition. (Hume 1978, 21)

Hume's view is not that we have a particular idea, or mental symbol, that is always used when thinking about a particular concept. What we have is a representation of a label that operates as a hub, linking together all the distributed representations of instances of the concept.

4.3. Rule following

The second form of systematicity that is supposed to be lacking in associationist systems is systematicity of inference or rule-following. As Fodor and Pylyshyn put it, "inferences that are of similar logical type ought, pretty generally to elicit correspondingly similar cognitive capacities" (1988, 43). The LoT hypothesis guarantees this systematicity by positing that thought is governed by syntactical and logical rules; the thought that Q follows from the thoughts that P and $P \rightarrow Q$ as a matter of cognitive architecture, no matter what P and Q refer to. Further, our rule-governed logical thought is supposed to be deductive and exact rather than probabilistic.

However, if a system does not come equipped with such rules, then there is no guarantee that it will be systematic in this way. For example, consider a connectionist network containing nodes labeled "A & B," "A," and "B." Because connectionist nodes become associated in virtue of experience rather than syntactic form, we can train the network such that "A & B" causes the "A" but not "B" node to activate (Fodor and Pylyshyn 1988). The argument is as follows:

1. If an associationist system makes a logical inference of form P, it is because of "statistical properties of the machine's training experience" (ibid., 28).
2. If a system has different statistical experiences for different instances of form P, then it will perform differently for different instances of the same logical form.

C: An associationist system's inferences will not be systematic.
C: An associationist system's rule adherence will be probabilistic.

For two reasons, Hume's first response is to argue that our thought is much less exact and systematic than alleged. First, he argues that even in our mathematical reasoning, the exactness of an idea cannot exceed the exactness of an impression; "the ideas which are most essential to geometry, *viz.* those of equality and inequality, of a right line and a plain surface, are far from being exact and determinate" (Hume 1978, 50–51). Second, our adherence to logical rules will be both probabilistic and context sensitive, for "the circumstance has a considerable influence on the understanding, and secretly changes the authority of the same argument, according to the different times, in which it is proposed to us" (ibid., 143).

Predictably, his next move is the externalist gambit. To the extent that our inferences *are* precise and systematic, it is because natural language is precise and

https://doi.org/10.1017/psa.2023.139 Published online by Cambridge University Press

systematic. On Hume's initial picture of language, we associate ideas (derived from the sight or sound) of words with ideas (mental pictures) of objects. However, he notes that sometimes we think purely using associations between words or symbols. These external tokens can have an exactness that surpasses that of our other sensory impressions:

> When we mention any great number, such as a thousand, the mind has generally no adequate idea of it, but only ... its adequate idea of the decimals, under which the number is comprehended. This imperfection, however, in our ideas, is never felt in our reasonings.[8] (ibid., 23)

To the extent that our mathematical or logical reasoning is systematic and precise, it is because it has been conjoined with (or indeed, performed entirely using) natural language symbols that are systematic and precise. By assigning numbers to the lengths of a triangle's sides, they come to stand in more precise relations to one another than they do in our impressions. Formal rules of geometry are reused across contexts, ensuring that our inferences about triangles come to be systematic.

5. Conclusion

Critics of associationism allege that it cannot account for the abstractness, systematicity, and precise rule-governed nature of human thought. In response, Hume argues that without language, our thought tends to be particular rather than abstract, context-sensitive rather than systematic, and probabilistic rather than exact. Thus, his account can explain those forms of thought that do not conform to linguistic or logical exactness. Further, if thinking with natural language tends to differ from thought without it, then a reasonable inference is that natural language is *causing* our thought to be abstract, systematic, and exact.

As I have shown, this basic picture has been adopted by bevy of researchers who see language as an external tool that can reshape our thought from the outside in. None of what I have said constitutes a *defense* of the externalist gambit. Regardless of its ultimate merits, there are two lessons to draw. First, the similarity between modern-day objections to connectionist architectures and the objections facing Hume's system is instructive. It shows that some problems raised for, say, DNNs, are targeting fundamental properties of associationist systems.

Hume's responses are equally instructive. That he gives a unified response to each of these three objections shows that the externalist gambit is not merely a piecemeal and accommodationist strategy for showing how such systems can instantiate human or human-like thought. For Hume, the three objections to his view are, at root, the same, and they call for the same response. The structure of our mind is derived from the structure of the world. If you want to explain how our thoughts can come to have linguistic structure, it is linguistic structure to which you should look.

Competing Interests. The author declares none.

[8] A similar point is made by Clark (2006, 297).

https://doi.org/10.1017/psa.2023.139 Published online by Cambridge University Press

References

Berkeley, George. 1975. A "Treatise Concerning the Principles of Human Knowledge." In *Berkeley: Philosophical Works*, edited by M. R. Ayers, 61–128. London: J. M. Dent & Sons.

Bradshaw, D. E. 1988. "Berkeley and Hume and Abstraction and Generalization." *History of Philosophy Quarterly* 5 (1):11–22.

Buckner, Cameron. 2018. "Empiricism without Magic: Transformational Abstraction in Deep Convolutional Neural Networks." *Synthese* 195 (12):5339–72. https://doi.org/10.1007/s11229-018-01949-1

Buckner, Cameron. 2019. "Deep Learning: A Philosophical Introduction." *Philosophy Compass* 14:1–19. https://doi.org/10.1111/phc3.12625

Buckner, Cameron. 2023. *From Deep Learning to Rational Machines: What the History of Philosophy Can Teach Us about the Future of Artificial Intelligence*. New York: Oxford University Press.

Clark, Andy. 2006. "Material Symbols." *Philosophical Psychology* 19 (3):291–307. https://doi.org/10.1080/09515080600689872

Clatterbuck, Hayley. 2016. "Darwin, Hume, Morgan, and the Vera Causae of Psychology." *Studies in History and Philosophy of Biological and Biomedical Sciences* 60:1–14. https://doi.org/10.1016/j.shpsc.2016.09.002

Fodor, Jerry, and Zenon Pylyshyn. 1988. "Connectionism and Cognitive Architecture: A Critical Analysis." *Cognition* 28 (1–2):3–71. https://doi.org/10.1016/0010-0277(88)90031-5

Hume, David. 1978. In *Treatise of Human Nature*, edited by L. A. Selby-Bigge and P. H. Nidditch. New York: Oxford University Press (first published 1738). https://doi.org/10.1093/oseo/instance.00046221

Inukai, Yumiko. 2010. "Hume on Relations: Are They Real?" *Canadian Journal of Philosophy* 40 (2):185–209. https://doi.org/10.1353/cjp.2010.0003

Lupyan, Gary, and Benjamin Bergen. 2016. "How Language Programs the Mind." *Topics in Cognitive Science* 8 (2):408 –24. https://doi.org/10.1111/tops.12155

Lupyan, Gary, David H. Rakison, and James L. McClelland. 2007. "Language Is Not Just for Talking: Redundant Labels Facilitate Learning of Novel Categories." *Psychological Science* 18 (12):1077–83. https://doi.org/10.1111/j.1467-9280.2007.02028.x

Machery, Edouard. 2005. "You Don't Know How You Think: Introspection and Language of Thought." *The British Journal for the Philosophy of Science* 56:469–85. https://doi.org/10.1093/bjps/axi130

Mandelbaum, Eric. 2020. "Associationist Theories of Thought," *The Stanford Encyclopedia of Philosophy* (Fall Edition), edited by Edward N. Zalta. Stanford: Stanford University Press. https://plato.stanford.edu/archives/fall2020/entries/associationist-thought/

Pylyshyn, Zenon. 1991. "Rules and Representations: Chomsky and Representational Realism." *The Chomskian Turn*, edited by Asa Kasher. Oxford: Basil Blackwell Limited.

Shea, Nicholas. 2007. "Content and Its Vehicles in Connectionist Systems." *Mind & Language* 22 (3):246–69. https://doi.org/10.1111/j.1468-0017.2007.00308.x

Silver, David et al. 2017. "Mastering the Game of Go without Human Knowledge." *Nature* 550 (7676): 354–59. https://doi.org/10.1038/nature24270

Smith, Linda B., and Diana Heise. 1992. "Perceptual Similarity and Conceptual Structure." In *Advances in Psychology* (Vol. 93, pp. 233–72). North-Holland. https://doi.org/10.1016/s0166-4115(08)61009-2

Waxman, Sandra R., and Dana B. Markow. 1995. "Words as Invitations to Form Categories: Evidence from 12-to 13-Month-Old Infants." *Cognitive Psychology* 29 (3):257–302. https://doi.org/10.1006/cogp.1995.1016

Cite this article: Clatterbuck, Hayley. 2024. "Hume's Externalist Gambit." *Philosophy of Science* 91 (5): 1296–1305. https://doi.org/10.1017/psa.2023.139