

**Davidson, Natural Selection and
the Principle of Charity**
John Clevenger
(*California State University, Bakersfield*)

In an earlier version of this paper, I defended Davidson's version of the Principle of Charity against a few very specific criticisms. My thoughts about the matter have changed much since then, and the current mongrelized instantiation is a direct result of this change. I still feel a strong inclination to defend Davidson's project, but *not* along characteristically Davidsonian lines. He put forth quasi-transcendental arguments intended to show that a being could not have thoughts (intended here mainly to mean propositional thoughts) unless that being had been embedded in a certain kind of community; namely, a community of similarly embodied creatures that interpret one another. I find this particular historical account of mind quite promising, though I have become quite skeptical that Davidson's arguments can get us to such a conclusion. Therefore, I offer in this paper another line of support for one of the claims Davidson's project must secure if it is to succeed: namely, that we have good reason to think that most of our perceptual beliefs about our environment are correct. As I will show, Davidson requires that we be able to assume that others perceive the environment in the same way that we do, which is to say, accurately. This claim, I will argue, is the most likely explanation for a large amount of empirical evidence. But first, I will sketch Davidson's account of radical interpretation, and show why the claim in question is vital to his project.

I. Radical Interpretation

Like many others, I wanted answers to such questions as 'What is meaning?', and became frustrated by the fatuity of the attempts at answers I found in Ogden and Richards, Charles Morris, Skinner and others. So I substituted another question which I thought might be less intractable: What would it suffice an interpreter to know in order to understand a speaker of an alien language, and how could he come to know it? (Davidson 1973)

The crucial insight here is that, if communication is indeed possible, then the full meaning of our speech must be transmittable through interpretation. In other words, there can be *nothing* more to the meaning of our words than what is exhausted in the process of interpretation. Thus, if we could rationally reconstruct what information we would need to know that would allow us to interpret a speaker, and how we could come to know it, then we would have given a general sketch of the empirical constraints that must be placed on any theory of meaning. What Davidson's project would show, if it were successful, is that the meaning of a word is not some abstract entity, but rather a complex relation between behavioral dispositions and the social and physical environments in which those dispositions are embedded. In fact, if Davidson is correct, we not only get a deflationary account of meaning, but a deflationary account of folk psychological ontology as well. Beliefs and desires are seen as theoretical compression algorithms initially invoked only when needed to organize and explain the behavioral dispositions of others. Davidson also takes the further step of arguing that, not only are propositional attitudes theoretical constructs, but we can only invoke them if we have been in contact with others. And without social interaction with other beings that we can interpret and be interpreted by, we would never be able to ascribe propositional attitudes to ourselves. In fact, without language, which is dependent upon social interpretation for Davidson, we wouldn't be able to have highly specific beliefs or desires at all. The basic point is thus: if we don't identify with other beings, then we don't have access to the concepts that allow propositional attitudes, propositional thought, and even an awareness of objectivity.

II. The Process

The thought experiment of the radical interpreter is meant to elucidate the empirical constraints on the nature of meaning. It's radical in the sense that the situation is one of two speakers possessing languages that have never come in contact with each other. That is, the speakers have no pre-existing way of translating each other's language. If it can be shown how, and through knowledge of what, these two speakers can come to understand one another, we should learn something interesting about the nature of meaning, and, if we buy into the rest of Davidson's project, about the nature of mind in general.

Davidson construes the project of radical interpretation to be, at a basic level, grounded in our ability to detect “hold-true” attitudes in others. This is simply the ability, deception aside, to tell if a speaker intends a certain sentence she utters to be true. Hold-true attitudes, he contends, are a function of sentence meaning and speaker belief. If a speaker utters, “That monkey is purple,” then the instances in which she holds that sentence true will be dependent upon the intended meaning of the sentence, and what she believes about her environment. The problem, though, is that even if we can identify that a speaker is holding a certain sentence true, we still have two variables left in our equation. First, when the speaker utters the sentence “That monkey is purple,” does she really mean that there’s a purple monkey over there, or does this sentence mean something different to her than it does to us? Second, does the speaker perceive the environment in the same way we do? Does she actually see that purple monkey over there? If we can’t find a way to hold one of these variables constant, we won’t be able to fix our interpretation of the speaker. Davidson’s solution to this problem is the Principle of Charity, to which we will return in a minute.

I think it’s pertinent here to mention an analogy to decision theory that heavily influenced Davidson. Once it’s articulated, the process of radical interpretation becomes much more intuitive. The methods of decision theory allow us to determine what an agent believes about the probability of certain events, and the subjective value they assign to possible outcomes, using solely their preferences about these events. Let’s say we have event X, where I accept a bet where I will receive 1000 dollars if the United States invades Iran and 2000 dollars if the United States doesn’t invade Iran. Now, given that we have just my willingness to accept that bet, we can’t derive what I believe about the probability of the United States invading Iran, or how I value my money. But, if we also observe event Y, where I accept a bet where I will receive 2000 dollars if the United States invades Iran and 1000 dollars if the United States doesn’t invade Iran, then we can derive that I believe the probability of the U.S. invading Iran to be 1/2. As Lepore and Ludwig note,

Thus, by identifying a special pattern among the agent’s preferences, one can solve for subjective probability by just using the data in the form, ultimately, of the agent’s choice behavior. After identifying the subjective probability the agent assigns to one sort of event, one can use it in turn to investigate his rankings among others on the basis of the relative desirability he assigns to them, and use this information in turn to investigate the subjective probability he assigns. (Lepore and Ludwig 2005, p. 181)

Davidson construes radical interpretation as similar to this task of assigning value and probability in decision theory. One uses an individual's behavior, their preference for a certain utterance, their hold-true attitude, to construct a hypothesis about the meaning of the utterance. The problem at this point, though, is that this hypothesis still has two variables left. In decision theory they are relative value and subjective probability. In radical interpretation they are sentence meaning and speaker belief. As we have seen, decision theory relies upon an ability to use the patterns that emerge in an agent's choices to fix subjective probability. Analogously, radical interpretation uses the patterns that emerge in a speaker's behavior to fix sentence meaning.

III. The Principle of Charity

The Principle of Charity, broadly construed, holds that we should interpret speakers to be as rational (read: consistent) and accurate as reasonably possible. More specifically, we should interpret a speaker as having mostly correct beliefs about their environment, and also generally consistent beliefs. What the Principle of Charity allows us is the capability to hold a speaker's beliefs about her environment constant, while we solve for her sentence meaning. When we interpret another's sentences we are, in essence, putting ourselves into her shoes. We assume that both of us perceive our shared environment in a similar way, and we assume that both of us operate according to the same general rational principles. If a speaker utters the word "flower" whenever she is pointing to a flower, we can, by assuming that she accurately perceives the structure of her environment, postulate that her utterance of the word "flower" actually means flower. We can then, through empirical testing, gather evidence to confirm or falsify our theory that the utterance "flower" really means flower. If all further evidence supports our theory, then we can conclude, with reasonable certainty, that the theory is correct. Thus, we have radically interpreted the speaker's language. But, we can only succeed in this project if we can hold one of the two variables mentioned above constant. For this, it seems, we need the Principle of Charity.

IV. A Bayesian Intuition Pump

In the second part of this paper I will develop the positive thesis that we have good reason to think that most of our perceptual beliefs about the world are true. I will start by arguing that the most likely explanation of some empirical observations is that organisms *can*

accurately perceive their environment. I will then argue that natural selection provides the means to explain how this is possible.

The history of evolutionary development on this planet has been one of struggle. Organisms compete with one another for a limited supply of those resources necessary for survival and reproduction. Along the way, they must avoid many potentially fatal pitfalls. Organisms must avoid predators and they must correctly identify edible food from poison. Higher-level organisms must also correctly identify potential mates for reproduction. If reproduction is successful, organisms must then use their discriminative abilities to provide not only for themselves, but also for their offspring. All of these tasks seem to be reliant upon an ability to accurately perceive the environment in which the organisms live. I use the word “seem” here because I want to stress the fact that these tasks are not *logically* dependent upon accurate perceiving the environment. It is possible, in some world, that an organism that possessed a perceptual system that systematically misled them about the structure and content of their environment could successfully complete these tasks. The odds of this happening, though, are so incredibly low that it shouldn’t be considered a viable option. And if the odds of a single organism braving a world of predators and pitfalls to successfully reproduce are statistically insignificant, then the odds of an entire lineage of organisms successfully managing to complete the tasks necessary for reproduction is astronomically low.

Let’s examine the following two hypotheses in a Bayesian context:

1. Organisms are systematically unable to accurately perceive their environments.
2. Organisms are able to accurately perceive their environments.

The key here is to construe the comparison of these two hypotheses as a likelihood inference. We should examine a set of observations and try to give a rough approximation of the probability of each observation given each of the competing hypotheses. Let’s assume that we observe an organism perform the following tasks:

1. Selectively choose between possible food sources, i.e. successfully gathering and eating enough food to provide adequate nutrition.
2. Recognize its offspring as such.
3. Recognize its mate as such.

4. Recognize a potential predator flying above, and successfully recognizing and navigating into a hole in the ground that provides safety.
5. Recognize lakes and streams as sources of water.
6. Recognize that water is different from other potentially harmful liquids.
7. Recognize that another member of its species is *not* a family member.

Now, consider the probability of the organism successfully completing each of these tasks in light of our two competing hypotheses. Also, consider the probability of the organism successfully completing *all* of these tasks in light of our two competing hypotheses. It's pretty clear that $\text{Pr}(\text{One of these observations} / \text{Hypothesis 2 (that the organism can accurately perceive their environment)})$ is greater than the $\text{Pr}(\text{One of these observations} / \text{Hypothesis 1 (that the organism cannot accurately perceive their environment)})$. **But** the $\text{Pr}(\text{All of these observations} / \text{Hypothesis 2})$ is exponentially greater than the $\text{Pr}(\text{All of these observations} / \text{Hypothesis 1})$. If this is correct, then the $\text{Pr}(\text{A lineage of organisms completing all of these tasks} / \text{Hypothesis 2})$ is so much greater than the $\text{Pr}(\text{A lineage of organisms completing all of these tasks} / \text{Hypothesis 1})$ that the probability of hypothesis 1 must be miniscule.

This does not mean that there is no evidence in support of hypothesis 1. There are obviously a number of observations that would, in a likelihood comparison, ostensibly favor hypothesis 1. If we observe our organism approach a predator, seemingly unaware of the danger, and subsequently become food, we may have reason to conclude that the organism's perception of the environment is faulty. And, of course, there are many more observations of this kind that might cause us to assign a greater probability to hypothesis 1. But, our hand is not forced in this direction. We can give plausible explanations for the observation without falling back on the general hypothesis that the organism is mostly incorrect about its environment. Perhaps our organism was distracted or injured in some way that impaired its vision. Perhaps the predator was camouflaged. We can give a number of plausible explanations for our organism's misstep that are consistent with hypothesis 2. Nonetheless, even if we grant that some observations will, in a likelihood comparison, support hypothesis 1, I believe that a holistic observation of the organism's behavior will still implicate hypothesis 2. The organism, if it is not a newborn, will have completed an infinite number of actions that point at hypothesis 2, *before* it is ever eaten by our predator. Thus, almost all observations that support

hypothesis 1 will be preceded by *many more* actions that point to hypothesis 2, and, if taken against the background of such actions, should be seen as either anomalies or an inability for our organism to accurately perceive the environment to the relevant level of detail. This last point is extremely important. Even if our organism does succumb to predation, it doesn't necessarily mean that it didn't accurately perceive its environment, it could be that its discriminative abilities simply could not accurately perceive those features of their environment that were relevant to staying alive. For Davidson, this is not a problem. All we need to secure the Principle of Charity is for humans to accurately perceive those features of the environment that other humans have access to. This is to say, we may not be able to accurately perceive infrared light, but we can all certainly accurately perceive red light.

V. The Argument From Natural Selection

The contention that most of our beliefs about the environment could be false is, I believe, based upon a failure to understand how natural selection operates. Given genetic variation, heredity, and selection (the differential success of organisms usually based on limited resources and/or competition), those organisms that do survive do so because they are better than their competitors at surviving. Over time, all organisms change, though those that change for the better (for our purposes those that perceive the environment accurately) will be more likely to survive. Of course it's possible that some organisms will survive because some of the conditions mentioned above are not met, or simply by dumb luck (remember the admission that none of this is logically necessary). There can be nothing logically necessary about it because the argument is dependent upon a number of contingent factors. Given an island with abundant food, no predators, and very little danger, there will not be such a pressure on the development of finely tuned perceptual abilities. In such a situation, an appeal to selection would be out of place. But, given the actual situation on this planet for the last few billion years, and given the physical evidence, the conclusion here seems pretty safe.

What I am trying to emphasize here is the fact that our perceptual abilities, and not simply our perceptions, are causally tied to the world. The development of sensory systems happens within a given environment, and the nature of this environment will causally affect how the systems are built. The evolutionary process of selection is a type of feedback loop. Organisms, in a sense, try out new phenotype characteristics through mutation. The structure and nature of the

organism's environment then determines how effective the new designs will be. If an organism is primarily nocturnal, the development of infrared vision, or sonar, will be highly adaptive, and the probability that the organism will survive, and pass on its improved perceptual system is increased.

This argument, with its focus upon perceptual accuracy, gets us a long way toward securing the Principle of Charity. But, it doesn't get us all the way there. We need two additional steps:

1. Perceptions have propositional content
2. *Ceteris Paribus*, we are justified in believing our perceptions.

If we can secure these, then it looks as though we've got our proverbial hooks into the world. I have already argued that our perceptual systems accurately portray the environment, what's needed now is reason to believe that our beliefs about the environment are related to our perceptual systems in such a way as to preserve that accuracy.

Humans obviously have beliefs, in whatever ontological manner one wants to characterize them, and some of these beliefs are caused by our perceptual systems. We don't require any justification, other than perceptual experience, for our perceptual beliefs. If I ask you why you believe that a certain dog is brown, and you respond that you perceived it to be brown under good lighting, then there isn't really much I can say. Sure, you could be wrong. You could be having vision problems. But, generally, your vision is good enough to warrant your belief. Davidson writes,

"The essential point is that we are justified in accepting the beliefs we are caused to have by the workings of our senses, not because our senses provide us with reasons or evidence, but because of the nature of perception. This is what I argued for when I said that perceptual beliefs are "veridical," meaning not that they are always true, but that we are justified in taking them to be true until shown wrong." (Davidson 1999, p. 254)

When Davidson speaks of the nature of perception, he is, perhaps unknowingly, relying upon the mechanism of evolution by natural selection. It is only in this context that we can give an account of why our perceptual abilities *should* generally be reliable. And it follows that, if we are generally justified in believing our perceptions because they're reliable, then most of our beliefs about the environment are probably true.

VI. Reflections

This conclusion is quite intuitive. When I point to a Great Dane and remark at how big it is, I expect our companion to have a similar perception. Perhaps she is more knowledgeable about Great Danes than I, and remarks that it's a rather small representative of the breed, but I certainly expect her to recognize it as a dog. I expect her to recognize that it has a certain mass, a certain color, and occupies a certain space. I expect her, to put it bluntly, to agree with me on most of my beliefs about the environment. I also expect that we are both mostly correct about the environment. And if our survival as a species really is dependent upon how well we're able to detect predators, find food, select mates and, in general, accurately perceive our environments, it seems outlandish to suggest that humans all inaccurately perceive the environment in the same way. It is logically possible, we shouldn't forget, that we could all be massively mistaken about the environment. But, taking the empirical fact that humans have survived such a long, intense, struggle for existence, there seems no way to claim this possibility is anything more than logical.

REFERENCES

- Davidson, D. (1973). "The Material Mind." In *Essays on Actions and Events*, pp. 245-260. New York: Clarendon Press.
- Davidson, D. (1999). "Reply to Burge." In L. Hahn (ed.), *The Philosophy of Donald Davidson, Library of Living Philosophers Volume XXVII*. Chicago and La Salle, Ill: Open Court.
- Lepore, E. and Ludwig, K. (2005). *Donald Davidson: Meaning, Truth, Language and Reality*. New York: Clarendon Press.