Aporia vol. 19 no. 1—2009

Hempel's Raven

Joshua Ernst

In his paper "Studies of Logical Confirmation," Carl Hempel discusses his criteria for an adequate theory of confirmation. In his discussion, he argues that Jean Nicod's criterion of confirmation fails to give an adequate theory because it generates a logical paradox, sometimes referred to as Hempel's Raven. Although some have argued that the paradox is moot for various reasons, Hempel argues that the solution given to the paradox is important. In this paper I will present Hempel's argument against Nicod, and the paradox that results under what Hempel calls "The Equivalence Condition." I will then give Hempel's own argument regarding the paradox and his proposed solution. Lastly, I will show an argument by I. J. Good that argues against the paradox and explain why Hempel's solution to the paradox has greater meaning for us.

Hempel first quotes from Nicod's "Foundation of Geometry and Induction:"

> Consider the formula or the law: *A entails B*. How can a particular proposition, or more briefly, a fact, affect its probability? If this fact consists of the presence of B in a case of A, it is favorable to the law "*A entails B*"; on the contrary, if it consists of the absence of B in a case of A, it is unfavorable to this law . . . . Thus, the entire influence of particular truths or facts on the probability of universal propositions or laws would operate by means of

Joshua Ernst is a senior majoring in philosophy with minors in logic and Spanish. His interests include philosophy of language, logic, and philosophy of religion. In the fall of 2009 he will be attending law school at Arizona State University. This essay took first place in the 2009 David H. Yarn Philosophical Essay Contest.

> these two elementary relations which we shall call *confirmation* and *invalidation*. (Hempel 9)

Nicod's hypothesis results in a universal hypothesis of the following form:

$$(x)(Ax \supset Bx)$$

Hempel argues, "In other words, an object confirms a universal conditional hypothesis if and only if it satisfies both the antecedent . . . and the consequent . . . of the conditional" (10). If I have an object that satisfies the antecedent, or is an instance of property A, say A*h*, and that same object satisfies the consequent, or has property B such that B*h*, then that object confirms the universal conditional hypothesis. An object then disconfirms the universal conditional hypothesis if it has both property A *and not* property B (A*h* and ~B*h*).

Hempel then examines the shortcomings of Nicod's theory, and subsequently establishes the paradox. Hempel gives two sentences, "All ravens are black" and "All non-black things are not ravens," as follows:

$$S_1: (x)(\text{Raven}(x) \supset \text{Black}(x))$$

$$S_2: (x)({\sim}\text{Black}(x) \supset {\sim}\text{Raven}(x))$$

To generate the problem, Hempel describes four different objects, *a*, *b*, *c*, and *d*, as follows: *a* is a raven and is black, *b* is a raven but is not black, *c* is not a raven but is black, and *d* is neither a raven nor black. "Then . . . *a* would confirm S_1, but be neutral with respect to S_2; *b* would disconfirm both S_1 and S_2; *c* would be neutral with respect to both S_1 and S_2, and *d* would confirm S_2, but be neutral with respect to S_1" (Hempel 11).[1] If we consider S_1 and S_2, the sentences say the same thing, they are merely written in a different way. And yet *a* and *d*, which satisfy S_1 and S_2 respectively, do not satisfy the other sentence. Thus, "Nicod's criterion makes confirmation depend not only on the content of the hypothesis, but also on its formulation" (Hempel 11). Hempel then arrives at a pivotal conclusion:

> Every hypothesis to which the criterion is applicable, i.e. every universal conditional, can be stated in a form for which there cannot possibly exist any confirming instances. Thus, the sentence
>
> $$(x)[(\text{Raven}(x) \cdot {\sim}\text{Black}(x)) \supset (\text{Raven}(x) \cdot {\sim}\text{Raven}(x))]$$

[1] For the purposes of this paper, I will not consider Hempel's concept of neutrality in a theory of confirmation.

> is readily recognized as equivalent to both S_1 and S_2 above; yet no object whatever can confirm this sentence. (Hempel 11)

From the conclusions reached above, Hempel can generate the paradox. He describes "a condition which an adequately defined concept of confirmation should satisfy . . . : [the] Equivalence condition: Whatever confirms (disconfirms) one of two equivalent sentences, also confirms (disconfirms) the other" (Hempel 12). Armed with the Equivalence condition (to be referred to as EC), Hempel shows the paradox that follows. Take our two sentences, S_1 and S_2. Everyone can agree, Hempel argues, that if *a* is "both a raven and black, then *a* certainly confirms S_1" (Hempel 14), and likewise, if *d* is neither black nor a raven, then it certainly confirms S_2. Using the EC, then, because S_1 and S_2 are equivalent, *d* is also confirming S_1. Thus any object which is not black and not a raven confirms our universal conditional hypothesis that all ravens are black. "Consequently, any red pencil, any green leaf, and yellow cow, etc., becomes confirming evidence for the hypothesis that all ravens are black" (Hempel 14). Hempel further defends this position:

> The following sentence can readily be shown to be equivalent to S_1: S_3: '$(x)[(\text{Raven}(x) \vee {\sim}\text{Raven}(x)) \supset ({\sim}\text{Raven}(x) \vee \text{Black}(x))]$', i.e. "Anything which is or is not a raven is either no raven or black." According to the above sufficient condition, S_3 is certainly confirmed by any object, say *e*, such that (1) *e* is or is not a raven and, in addition, (2) *e* is not a raven or also black. Since (1) is analytic, these conditions reduce to (2). By virtue of the equivalence condition, we have therefore to consider as confirming for S_1 any object which is either no raven or also black (in other words: any object which is no raven at all, or a black raven). (Hempel 14)

Hempel wonders if the customary mode of presenting

> sentences of universal conditional form . . . has to be modified; and perhaps such a modification would automatically remove the paradoxes of confirmation[.] If this is not so, there seems to be only one alternative left, namely to show that the impression of the paradoxical character of those consequences is due to misunderstanding and can be dispelled, so that no theoretical difficulty remains. (Hempel 15)

I will focus on the second of these two possibilities, "the possibility of tracing the impression of paradoxicality back to a misunderstanding" (Hempel 15).

In defense of the second of the two possibilities, Hempel argues that the conclusion is actually a psychological illusion rather than something that is objectively founded. He gives two arguments in support of this theory. The first comes in how we understand a universal conditional hypothesis, like $(x)(Px \supset Qx)$. In saying that every *P* is *Q*, we think that we are limiting our focus to the objects that have the property *P*. Hempel says, "This idea involves a confusion of logical and practical considerations" (Hempel 18). While we may have only written about the property *P* in our sentence, that universal conditional hypothesis asserts something about, "and indeed imposes restrictions upon," all objects (Hempel 18). The statement "every *P* is *Q*" does not allow an occurrence of an object with predicate *P* to occur *unless* it is also has property Q: "Every object either belongs to this class or falls outside it, and thus, every object—and not only the P's—either conforms to the hypothesis or violates it; there is no object which is not implicitly 'referred to' by a hypothesis of this type" (Hempel 18–19).

Hempel gives the second argument for his hypothesis in the following example:

> Suppose we have a universal conditional hypothesis which states "All sodium salts burn yellow." Someone comes with a piece of ice and holds it in a colorless flame. As is expected, the flame would not turn yellow, and this would support the hypothesis that "Whatever does not burn yellow is not sodium salt." Suppose another person comes with an object whose "chemical constitution is as yet unknown to us" and when held under the flame "fails to turn it yellow, and where subsequent analysis reveals it to contain no sodium salt." (Hempel 20)

We would say that this experiment also confirms our hypothesis that all sodium salts burn yellow and that the two experiments confirm the same thing, "no matter in which of its various equivalent formulations it may be expressed; thus, the data here obtained constitute confirming evidence for the hypothesis" (Hempel 19). The paradox comes when we consider a fundamental difference between the two experiments. In the first test, we knew that the object being held in the flame was ice, and because we had a previous knowledge that ice does not contain sodium salt, "the outcome of the flame-color test becomes entirely irrelevant for the confirmation of the hypothesis and thus can yield no new evidence for us" (Hempel 19).

Hempel argues that when a paradoxical situation arises, like the one demonstrated with the raven, we are not merely judging the relation of an object that is evidence of a claim to the claim itself. Instead,

> we tacitly introduce a comparison of [the hypothesis] with a body of evidence which consists of [the object] in conjunction with an additional amount of information which we happen to have at our disposal; in our illustration, this information includes the knowledge (1) that the substance used in the experiment is ice, and (2) that ice contains no sodium salt. If we assume this additional information as given, then, of course, the outcome of the experiment can add no strength to the hypothesis under consideration. (Hempel 19–20)

In order to avoid this "tacit reference to additional knowledge," we need to ask an important question: If I am presented with some object *a*, (assuming *a* to be a piece of ice and assuming this information is withheld from me), and when I hold *a* in a colorless flame which does not change while burning *a*, and thus *a* does not contain sodium salt, "does *a* then constitute confirming evidence for the hypothesis?" (Hempel 20). It is only when we do this that finding confirming evidence in support of a universal conditional hypothesis can occur, and the paradox vanishes.

Let us take our evidence and consider once again the raven's paradox. As stated above, the paradox arose when we were able to use an object which satisfies S_2, namely that all non-black things are not ravens, as evidence in support of the fact that all ravens are black. Suppose we consider the white computer that I am using to write this paper. This constitutes an object that is not black and not a raven. Thus my computer lends support to the fact that all ravens are black. However, suppose that an alien comes from another galaxy. This being has no idea what it means to be a raven, and in his world, in some unfathomable way to us, there is no experience of color. I tell him that I have a universal conditional hypothesis that I want him to confirm, namely that all ravens are black. The alien is able to look into my mind and extract a colorless representation of a raven. I then hold up a piece of black paper in order to allow him to understand the color black. Once he understands the hypothesis, I then show him my white computer. He can clearly see that the object does not match the mental representation he has of a raven, nor does the color of my computer match the color of the piece of paper I have shown him. He takes this as evidence that all non-black things are not ravens, and only in an instance like this, where we have no previous knowledge associated with the objects in question, will our results constitute confirming evidence, and as Hempel argued, the paradox vanishes.

In his article "The White Shoe Is a Red Herring," I. J. Good makes an argument against the raven's paradox by arguing that Nicod's criterion is false. Good gives an example in which finding a black raven would *decrease* the probability that all ravens are black. He argues that

> in the present note we show in six sentences, and even without reference to the contrapositive, that a case of a hypothesis does not necessarily support it.
>
> Suppose that we know we are in one or other of two worlds, and the hypothesis, H, under consideration is that all the crows in our world are black. We know in advance that in one world there are a hundred black crows, no crows that are not black, and a million other birds; and that in the other world there are a thousand black crows, one white one, and a million other birds. A bird is selected equiprobably at random from all the birds in our world. It turns out to be a black crow. This is strong evidence . . . that we are in the second world, wherein not all crows are black. Thus the observation of a black crow, in the circumstances described, undermines the hypothesis that all the crows in our world are black. Thus the initial premise of the paradox of confirmation is false, and no reference to the contrapositive is required. (Good 322)

Because finding an object that would normally support the universal conditional hypothesis in fact decreases the likelihood of that statement being true, then for a white shoe to give evidence to the fact that all ravens are black is not as important as Hempel believes it to be. However, if we consider Hempel's argument which I have just illustrated, I believe that Good's reason to ignore the paradox does not hold as strongly as he believes.

Good presents a situation which decreases the importance of an object giving support to a completely unrelated hypothesis by showing it is possible to have an object which confirms the hypothesis and actually *decrease* the probability of the hypothesis being true. However, I argue that Hempel's argument regarding the paradox has a greater bearing on our everyday life. Because we store information in a categorized way, barring some type of mental defect, the associations we make in our daily observations will stay with us. However, it is that very idea—that we make and keep sensory observations and that there are certain associations among the observations (all ravens are black)—that causes paradoxical situations to appear. The result is a contradiction between our senses and our reason.

The contradiction between reason and senses is not new. Zeno is credited with famous, ancient paradoxes which illustrate this very idea. Suppose we look at an arrow in flight. When we talk about its motion, we can either say that it moves in the space it is in or that it moves in the space in which it is not. The arrow cannot move in the space that it is in, for it occupies all of the space that is possible for the arrow to occupy. It also cannot move in a space in which it is not, for how can something move in a space which it does not occupy? Thus the arrow cannot actually move, and yet, our senses tell us that the arrow is moving. Oftentimes we are more inclined to trust our senses than our reason, so many would see the above paradox as merely twisting words. However, in the example of Hempel's raven, we are not twisting words nor committing any hidden fallacies, and yet the paradox arises. Hempel's explanation of the psychological illusion best answers the paradox, because it gives us a tool to begin to resolve the contradictions we see between reason and the senses. If we can disassociate ourselves from our observations, thus allowing reason and logic to properly function, then the paradoxes that we seem to observe will disappear.

Works Cited

Good, I. J. "The White Shoe is a Red Herring." *The British Journal for the Philosophy of Science* 17 (1967): 322.

Hempel, Carl G. "Studies in the Logic of Confirmation (I.)." *Mind* 54.213 (1945): 1–26.