A NOTE ON THE PARADOXES OF CONFIRMATION.

IN MIND for April 1945, Mr. C. H. Whiteley discusses certain peculiar features of the concept of confirmation, which were pointed out and analysed in my article, " Studies in the Logic of Confirmation " in the January and April issues of MIND for 1945, especially on pages 13-21. The peculiarities in question, which I called paradoxes of confirmation, pertain to the non-graduated relation of being-confirming-evidence-for a hypothesis, as contradistinguished from the concept of degree of confirmation of a hypothesis relatively to given empirical evidence. The paradoxes result from the following consideration: It seems reasonable to say that any proposed analysis or rational reconstruction of the non-graduated relation of confirmation, in order to be adequate, has to satisfy the following two requirements:

R1: Whenever an object has two attributes C_1, C_2, it constitutes confirming evidence for the hypothesis that every object which has the attribute C_1 also has the attribute C_2.

R2: An object which constitutes confirming evidence for one of two logically equivalent sentences, is confirming evidence also for the other (for the two sentences in this case express the same hypothesis).

Now R2 entails in particular that the following three sentences, being logically equivalent, have to be confirmed by the same class of objects:

S1: Whatever is a raven is black,
S2: Whatever is not black is not a raven,
S3: Whatever is or is not a raven is either no raven or black.

But according to R1, S3 is confirmed by any object which has the attributes (C_1) of being or not being a raven, and (C_2) of being no raven or black. As C_1 is analytic, it follows that each of the three sentences must be confirmed by any object that is either no raven or black. Thus, *e.g.*, to take Mr. Whiteley's illustrations, a red herring or a white elephant would constitute confirming evidence for S1. The general principle here illustrated together with its various consequences, which at first blush appear highly counter-intuitive, constitute the paradoxes of confirmation. In my article I tried to show that upon closer analysis the results thus arrived at prove to be reasonable, and that the impression of paradoxicality arises from a misguided intuition in the matter; the details of this argument, however, are not needed for a discussion of Mr. Whiteley's note.

Mr. Whiteley states that the results in question are "not only paradoxical, but false". If this is so, then obviously at least one of the premises R1, R2, from which they were derived, must be false, and Mr. Whiteley argues indeed that R1 does not generally hold. In elaborating this point, he argues that a distinction has to be made between the relation of being-evidence-for a hypothesis, which he claims does not satisfy R1, and the relation of being-an-instance-of an hypothesis; and while he does not provide a definition for this second relation, he seems to assume that at any rate it does satisfy R1 (when this condition is formulated for "instance of" instead of for "evidence for"), for he illustrates the instance relation by pointing out that any black raven is an instance of S1, any non-black non-raven an instance of S2, and anything that is not a non-black raven an instance of S3. This instance relation, Mr. Whiteley argues, is entirely different from the relation of being-evidence-for a hypothesis. Thus, *e.g.*, one instance of S1 does not constitute confirming evidence for S1 because "the occurrence of one black raven is not improbable even if S1 is false, and therefore provides no evidence for it"; but a series of black ravens, he states, does constitute confirming evidence because "If S1 is false, *i.e.*, if there are ravens which are not black, it is *a priori* unlikely that a collection of ravens taken at random will consist entirely of black ones". In connection with his assertion, finally, that the confirming-evidence relation does not generally satisfy R1, Mr. Whiteley states that while "S1 is established by observing instances of it", this is not true of S2, because "in S1 the conditions specified are positive, and in S2 they are negative"; and he proceeds to argue, in effect, that the instances of S2, whether taken singly or in collections, are irrelevant for the confirmation of S2.

Mr. Whiteley's suggestions do not seem to me to provide a basis for an adequate analysis of the paradoxes. My principal reasons for this opinion are as follows:

1. Mr. Whiteley does not give a definition or any other general characterisation of exactly what is to be understood by an instance of a hypothesis, although this concept is crucial for his argument. Let us assume however—and this seems to be a fair interpretation of his intentions—that he would call an object x an instance of a general hypothesis asserting that everything with the attribute C_1 also has the attribute C_2 if, and only if, x has both the attribute C_1 and the attribute C_2. This conception of the instance relation involves a serious difficulty: according to it, no instance of the sentence S1 above is an instance of S2, and some of the instances of S3 (namely those which are black non-ravens) are instances of neither S1 nor S2—despite the fact that all three sentences, being logically equivalent, express the same hypothesis. In fact, after having characterised the three different classes of objects which constitute instances of S1, S2, S3, respectively, Mr. Whiteley himself briefly

remarks that " in this respect the three sentences are not equivalent ", without, however, making any further comment on what is meant here by equivalence, or how the difficulty is to be met that sentences which make the same assertion differ in their instances.

2. These paradoxes of the instance relation also seem hardly reconcilable with Mr. Whiteley's assertion that " in order to know whether a particular fact is an instance of . . . a general hypothesis all that I need to know is the fact and the hypothesis " ; for, as the preceding illustration shows, one would also have to know by means of what particular sentence the hypothesis happens to be formulated ; and the class of instances of a given hypothesis would, in general, change with its formulation.

3. Besides, the general concept of instance, which seems intuitively so clear and unproblematic as hardly to require any explanation, actually involves very serious difficulties when applied to general hypotheses which contain relation terms. This can be seen from an argument on page 13, footnote, of my article, where it is shown that if for an ordered couple of objects, (a, b), both $\mathrm{R}(a, b)$ and $\sim \mathrm{R}(b, a)$ is the case, then the couple satisfies the two conditions $\mathrm{C}_1(x, y) : \sim (\mathrm{R}(x, y) \,.\, \mathrm{R}(y, x))$ and $\mathrm{C}_2(x, y) : (\mathrm{R}(x, y) \,.\, \sim \mathrm{R}(y, x))$; therefore, according to the intuitive notion of instance, any such couple (a, b) would constitute an instance of the following universal hypothesis : " Any couple which has the attribute C_1 also has the attribute C_2 " ; *i.e.*, ' $(x)(y)(\mathrm{C}_1(x, y) \supset \mathrm{C}_2(x, y))$ '. Actually, however, this hypothesis can be shown to be logically incompatible with ' $\sim \mathrm{R}(b, a)$ ', *i.e.*, with one of the two statements which seem to make (a, b) an instance of the hypothesis.

In this context, I should like to correct a misprint in the footnote on page 13 of my article : S1 should read as follows :

$$(x)(y)[(\sim R(x, y) \,.\, R(y, x)) \supset (R(x, y) \,.\, \sim R(y, x))].$$

4. The argument which Mr. Whiteley uses to show that one black raven cannot constitute confirming evidence for S1, while a series of ravens can, makes use of a number of concepts such as " probable ", " improbable ", and " *a priori* unlikely ", which have been employed in philosophical discussions in a variety of different meanings. Mr. Whiteley does not indicate which of these meanings he has in mind, and I am not aware of any theory of probability which could serve as a basis for the establishment of the argument which Mr. Whiteley presents in this context. Indeed, I doubt that there is such a theory ; for clearly, the difference between one instance and a series of instances is one of degree only, and it seems that, therefore, in any adequate theory of probability in which several instances may confirm a hypothesis—*i.e.*, in Mr. Whiteley's terms, " make it more or less probable "—a single instance must be able to do the same. (The degree of confirmation might, of course, be different, but that question is not involved here.)

5. The distinction between positive and negative conditions, which is presupposed but not explained in Mr. Whiteley's argument, appears to be quite problematic. In fact, the intended differentiation seems to be applicable to *predicate expressions*, which designate conditions, attributes, or properties, but *not to their designata:* It is obviously possible to distinguish between predicate expressions which begin with a denial sign and those which do not; but the same condition or attribute may be referred to by positive as well as by negative expressions in this sense. Thus, *e.g.*, in our example, it would be possible to introduce two special terms denoting non-black things and non-ravens respectively; and when formulated by means of these, the conditions specified in S2 would be "positive" and those in S1 "negative". Any argument which presupposes a distinction of positive and negative conditions appears, therefore, to be untenable.

CARL G. HEMPEL.