

## PHENOMENAL SIMILARITY

SYDNEY SHOEMAKER  
Cornell University

### I

I shall begin by quoting from some recent philosophical discussions of sensation. In his book *Myself and Others* Don Locke says that "My elders and betters taught me what 'pain' meant even though they did not and could not know how my sensations felt because they did not and could not feel them".<sup>1</sup> This is possible, according to Locke, because "A sensation's being a pain sensation is not a matter of how it feels, but a matter of its being of the sort caused by bodily damage and leading to pain behavior. Similarly a sensation's being a sensation of cold is not a matter of how it feels, but a matter of its being a sensation of the sort caused by frost and snow and leading to shivering, etc. And similarly for other sensations".<sup>2</sup> Alan Donagan expresses much the same view in his paper "Wittgenstein on Sensation", and attributes it to Wittgenstein. On this view, "you and I correctly say that we have the same sensation, say toothache, if we both have something frightful that we would naturally express by holding and rubbing our jaws, by certain kinds of grimace, and the like. Whether the internal character of what is expressed in these ways is the same for you as for me is irrelevant to the meaning of the word 'toothache'".<sup>3</sup> This view does not deny that there is and must be an inner "ac-

<sup>1</sup> Oxford, 1968, p. 98.

<sup>2</sup> *Op. cit.*, p. 101.

<sup>3</sup> In Pitcher, ed., Wittgenstein, *The Philosophical Investigations* (Garden City, N. Y., 1966), p. 348.

companionment” to pain behavior, having some “internal character” or other, if there is to be pain. “What is irrelevant”, says Donagan, “is not the existence of the object, but what it happens to be. You and I could not have a common word for pain unless our natural pain behavior was accompanied by something frightful; but whether the accompaniment is the same for both of us, or even whether it changes or not (provided we do not notice it) is irrelevant).”<sup>4</sup>

I assume that this view will strike nearly everyone, even including those who are persuaded that there is something to it, as *prima facie* paradoxical and counter-intuitive. It seems to imply that it is possible that a sensation that feels exactly like my most recent pain might be, in someone else or me at another time, a tickle rather than a pain, and that a sensation having the “internal character” of my most recent tickle might be a pain rather than a tickle. This conflicts with the intuition which Saul Kripke recently expressed by saying that pain “is not picked out by one of its accidental properties; rather it is picked out by the property of being pain itself, by its immediate phenomenological quality. Thus pain . . . is not only rigidly designated by ‘pain’ but the reference of the designator is determined by an essential property of the referent”.<sup>5</sup> I would suppose that what Kripke calls the “immediate phenomenological quality” of pain is what Donagan calls its “internal character” and what Locke refers to by speaking of “how it feels”. Thus Kripke affirms, what Locke and Donagan deny, that what constitutes something’s being a pain is its having a certain “internal character” or “phenomenological quality”. If there is a common-sense view on this matter, it is surely much closer to Kripke’s than to Locke’s and Donagan’s.

I suspect that the view of Locke and Donagan is the result of the pressure of two *prima facie* conflicting ideas. One of

<sup>4</sup> *Op. cit.*, pp. 346-347.

<sup>5</sup> “Naming and Necessity”, in Davidson and Harman, eds., *Semantics of Natural Language* (Dordrecht, Holland, 1972), p. 340.

these is the idea that the concept of a particular sort of mental state centrally involves, and perhaps can be defined in terms of, the concepts of certain of its typical behavioral effects or manifestations, and in some cases, the concepts of its bodily causes. This is an idea shared, in one version or another, by philosophical behaviorists, Wittgensteinians, and proponents of "causal" or "functional" accounts of mental states, including some proponents of Central State Materialism. In *prima facie* conflict with this is the idea—or I should say the fact—that each of us is directly aware of phenomenal similarities and differences between his own mental states, especially his sensory states, in a way that does not involve knowing about corresponding similarities and differences in the causes and causal powers of these states, and in a way that seemingly leaves it a contingent matter whether states that are given as similar have correspondingly similar causes and tendencies to influence behavior and whether states that are given as different have correspondingly different causes and tendencies to influence behavior. These phenomenal similarities and differences are thought to exist in virtue of the intrinsic features of the mental states in question, their "internal" or "phenomenal" character. Thus the connection between these intrinsic features of mental states, on the one hand, and bodily states of affairs, on the other, seems purely contingent. And if one combines this view with the view that our concepts of mental states centrally involve the concepts of the bodily causes and/or behavioral effects or manifestations of these states, e.g., the view of David Armstrong that a particular mental state can be defined as a "state apt for bringing about" certain behavior, one gets the view, held by Locke and Donagan, that while sensations and the like have phenomenal features, which we can be immediately aware of in introspection, these features are logically irrelevant to their belonging to the kinds marked out by our mental concepts, e.g., the concept of pain.

The notorious problem of the "inverted spectrum" is relevant here. It is often suggested that it might be, for all that can be shown by any possible behavior (including linguistic behavior), that blue things standardly look to one person the way yellow things standardly look to another, and likewise for other pairs of colors. Now when this alleged possibility is used to raise epistemological problems, the relevant phenomenal similarities and differences are envisaged as holding interpersonally, e.g., between my experiences and yours, and so as not being open to anyone's introspection. But I think that what primarily gives rise to the idea that such similarities and differences are possible interpersonally is the fact that we can imagine them occurring intrapersonally, where they would be immediately accessible to introspection or introspection *cum* memory. For example, it seems imaginable that my spectrum might become inverted relative to what it had been in the past so that after a certain time blue things look to me the way yellow things looked to me before that time, and so on. Not only can I imagine this happening to me; I can also imagine behavior (verbal and otherwise) that would be evidence that it had happened to someone else. But presumably someone to whom this happened would eventually adjust to the change, and accommodate his usage of color terms to that of other persons; once this accommodation has taken place he will say not only that the sky *is* blue but that it *looks* blue, despite the fact that it looks to him the way yellow things looked to him before the change. And here something like the Locke-Donagan description seems appropriate; if some underwent spectrum inversion, and accommodated to the change, it seems that we would say that at different times in his life experiences "of the same sort", e.g., experiences of blue, were phenomenally different, and sensations "of different sorts", e.g., an experience of yellow and one of blue, were phenomenally similar. And if interpersonal spectrum inversion is possible, this will amount to there being interpersonal differences between experiences of

“the same sort” and interpersonal similarities between experiences of “different sorts”, even if the persons having the experiences are behaving in identical ways in identical environments.

When we move from the case of visual experience to that of sensations like pain, the notion of “qualia inversion” becomes much more problematical. It is far from clear that we can make sense of the notion of someone having a sensation phenomenally just like pain which he does not find unpleasant or distressing, and which he responds to in the way other people respond to tickling sensations. Because this raises rather special problems, I shall say no more about pains and tickles and the like in this paper. I hope that what I have to say about phenomenal similarity can eventually be extended to cover the case of these so-called “intransitive” sensations. But my immediate purpose is to throw light on the application of the notion of similarity to perceptual experiences — for example, to similarities between the way various things look to us, or, in Chisholm’s terminology, between various ways of being “appeared to”.

## II

So much by way of introduction. I now want to consider the notion of similarity, first in general, and then as it applies to experiences. I shall approach this notion *via* a topic that may initially seem to have little if anything to do with it.

In his book, *Fact, Fiction and Forecast*, Nelson Goodman posed what he called “the new riddle of induction”. He introduced a predicate, “grue”, which is defined as applying “to all things examined before  $t$  just in case they are green but to other things just in case they are blue”, where  $t$  is some particular time, say Noon, Eastern Daylight Time, on July 4, 1984. Goodman’s definition has been interpreted in different ways by different commentators and critics. But for our present purposes it does not matter which interpretation

we adopt. On any of them it will be true that all emeralds examined before  $t$  were green when examined before  $t$  if and only if all emeralds examined before  $t$  were grue when examined before  $t$ , and on any of them it will be true that an emerald examined for the first time after  $t$  will be grue when first examined just in case it will be blue when first examined. And this is enough to generate Goodman's puzzle. Supposing that the time is before  $t$ , the fact that all emeralds so far examined have been green seems clearly to provide good inductive support for the generalization that all emeralds are green and thus for the conclusion that emeralds examined for the first time after  $t$  will be green. But using the same principles of inductive inference, the fact that all emeralds so far examined have been grue seems to provide equally strong inductive support for the generalization that all emeralds are grue, and thus for the conclusion that emeralds examined for the first time after  $t$  will be grue, from which it follows that they will be blue. Yet in a sense the fact that all examined emeralds have been grue seems to be the same as the fact that all examined emeralds have been green, given that the time is before  $t$ . Clearly we cannot allow that incompatible conclusions can legitimately be inferred from the same evidence; so the problem is to explain why one of these inductive extrapolations (presumably that involving "green") is legitimate and the other not.

There is another problem that is raised by predicates like "grue", although not only by such predicates. If A and B are both grue, they have something in common, namely being grue. But intuitively this doesn't guarantee that A and B are in any significant way similar; for it may be, compatibly with this, that A is green and B is blue, and that they differ in size, shape, composition, etc. Evidently it is not true of just any predicate that the things to which it applies are alike, or similar, in virtue of the facts that make it true that the predicate applies to both. Indeed, it is obvious on slight reflection that given any two things, no matter how dissimilar

they are (intuitively speaking), we can define any number of predicates that are true of both. Since sharing of predicates does not, in itself, give us similarity, what does give it to us? Another way of raising the same question is to ask what makes some classes, and not others, "natural kinds".

But now let us return to Goodman's "new riddle of induction". I shall not discuss here the solution Goodman offers to this "riddle", nor shall I discuss any other proposed solutions to it.<sup>6</sup> My present interest is not in the riddle as such but in a connection which Goodman implicitly establishes in posing it.

We of course believe that certain inductions involving the predicate "green" are acceptable while parallel inductions involving the predicate "grue" are not. In Goodman's terminology we can express this belief by saying that certain hypotheses involving "green" are "projectible" while corresponding hypotheses involving "grue" are not. Following the precedent set by other writers, including Quine, I shall depart from Goodman's usage by sometimes applying the terms "projectible" and "unprojectible" to predicates as well as to hypotheses containing predicates; thus I will say that the predicate "green" is projectible and that the predicate "grue" is unprojectible. Projectible predicates are those that are, in Goodman's words, "well behaved" in inductive inference; roughly speaking, they are such that generalizations involving them that have been observed to be true in examined cases can reasonably be believed to hold in unexamined cases as well. Where a predicate is projectible, I shall allow myself to speak of the corresponding *property* as projectible as well; so, for example, I will say that the property green is projectible. Needless to say, this is not Goodman's usage.

Now as Goodman points out, the distinction between the projectible and the unprojectible is closely connected with that between "lawlike" and "accidental" generalizations. Law-

<sup>6</sup> I present my views on this in "On Projecting the Unprojectible", *The Philosophical Review*, LXXXIV, 2 (April, 1975).

like generalizations, those that “sustain” counterfactuals, are confirmed by positive instances, and this will not in general be true of generalizations of the form “All A are B” unless either the predicates “A” and “B” are projectible or the generalization is equivalent to one of that form in which the “A” and “B” are projectible.

But the connection that especially interests me is that between the notions of projectibility and lawlikeness, on the one hand, and that of likeness or similarity, on the other. Goodman remarks that “the entrenchment of classes is some measure of their genuineness as kinds; roughly speaking, two things are the more akin according as there is a more specific and better entrenched predicate that applies to both”.<sup>7</sup> Here Goodman is assuming the truth of his theory about projectibility, in which the notion of projectibility is defined in terms of the notion of “entrenchment”; to oversimplify, a predicate is entrenched, and so presumptively projectible according to Goodman, to the extent that it, or other predicates having the same extension, have been projected in the past. But it seems clear that Goodman is asserting here a connection, what I would call a conceptual connection, between a predicate’s being projectible and its extension’s being a “natural” or “genuine” kind. The same connection is more explicitly asserted by Quine, who remarks that “a projectible predicate is one that is true of all and only things of a kind”.<sup>8</sup> And Quine goes on to link the notion of a kind with that of similarity or resemblance: “The notion of a kind and the notion of similarity or resemblance seem to be variants or adaptations of a single notion.”<sup>9</sup> Although talk of properties is not congenial to Goodman and Quine, given their extensionalism, it would also seem that our intuitive notion of a genuine property is closely linked with the notion of projectibility; we

<sup>7</sup> Goodman, *Fact, Fiction and Forecast*, 2nd edn. (New York, 1965), p. 121.

<sup>8</sup> W. V. O. Quine, “Natural Kinds”, in *Ontological Relativity and Other Essays* (New York, 1969), p. 116.

<sup>9</sup> *Op. cit.*, p. 117.



are reluctant to say that things that are grue thereby share a genuine property, and this is of a piece with our unwillingness to say that the extension of "grue" is a natural kind, or that things that are grue thereby resemble one another —and all of this is connected with our belief that "grue" is not a projectible predicate.

Neither Goodman nor Quine thinks that the notion of similarity provides any help in explaining or applying the notion of inductive projectibility, but in both writers there is, as we have seen, the suggestion that the notion of projectibility is of some use in explaining the notion of similarity. And in a recent essay entitled "Seven Strictures on Similarity" Goodman expresses the suspicion that "rather than similarity providing any guidelines for inductive practice, inductive practice may provide the basis of some canons of similarity".<sup>10</sup> On the rather slender basis of the passages I have quoted, I shall use the term "Goodman-Quine account of similarity" to refer to the view that the notion of similarity is to be explicated, in part, in terms of the notion of inductive projectibility. This name should not be taken too seriously; for in fact, Goodman and Quine hold the notion of similarity in such low regard that they probably would not want to dignify it by offering their remarks as an "account" of it. Quine speaks of the "dubious scientific standing" of this notion, and remarks that "it is a mark of maturity in a branch of science that the notion of similarity or kind finally dissolves, so far as it is relevant to that branch of science". Goodman uses even harsher terms; in "Seven Strictures on Similarity" he says that similarity is "a pretender, an imposter, a quack".<sup>11</sup> Still, Goodman and Quine seem to concede that there is *some* basis to the intuitive distinction between genuine similarity and the mere sharing of predicates, and their remarks suggest the view that this distinction coincides, at least roughly, with the

<sup>10</sup> In Foster and Swanson, eds., *Experience and Theory* (Amherst, Mass., 1970), p. 24.

<sup>11</sup> *Op. cit.*, p. 19.

distinction between projectible predicates and unprojectible ones.<sup>12</sup>

### III

Now there is, I think, a very natural resistance to the idea that the notion of projectibility must enter into a satisfactory explication of such notions as that of a natural kind, that of a genuine property, and especially that of similarity or resemblance. Similarity and dissimilarity seem to be paradigm cases of relationships that can be directly "given" in experience, that is, relationships we can directly observe to hold. The notion of projectibility, connected as it is with the notion of lawlikeness, seems a close cousin of the notion of causality. And due in part to the influence of Hume there is a tendency to think that no fact involving concepts in this family can be directly experienced.

I think myself that this Humean view is the source of a good deal of philosophical mischief. And while Goodman and Quine write as if Hume were their ally on the topic of induction, the implications of their views about similarity and natural kinds seem to me fundamentally at odds with Humean assumptions. For Hume never questioned our ability to perceive genuine similarities or resemblances, yet, notoriously, he regarded the projectibility of these similarities, and of the

<sup>12</sup> I should perhaps indicate that my agreement with the Goodman-Quine account is not as complete as the body of this paper may suggest. What is wrong with this account is not what it says but what it leaves out. It makes no mention of the notion of causality; yet it is *via* this notion, I think, that the notion of inductive projectibility is linked with the notions of natural kinds, genuine similarity, and genuine properties. I believe that the identity conditions for genuine properties, those the sharing of which contributes to genuine similarity, can be stated in causal terms; roughly, the identity of such a property consists in its potential for contributing to the causal powers of the things that have it. And it stands to reason that it will be predicates that stand for such properties that will be inductively projectible. Of course, Goodman's philosophical scruples do not permit him to rely on such notions as causality and potentiality in philosophical analysis (see FFF, p. 34), and the same is presumably true of Quine. But I think that these scruples are misguided.

properties in virtue of which they hold, as an open question which no sort of argument or evidence can close. Certainly Hume could not have raised the problem of induction in the way he does if he had thought that in being aware of similarities one is, *ipso facto*, aware of the sharing of projectible properties; for to hold the latter is to hold that in being aware of similarities one is already well on one's way towards knowing what inductions are legitimate.<sup>13</sup>

Still, the Humean view does have considerable initial plausibility, and if we are to defend what I am calling (with tongue in cheek) the Goodman-Quine account of similarity—the view that the notion of similarity is to be explained, in part, in terms of the notion of inductive projectibility—we must somehow explain away the seeming incompatibility between this account and the fact that similarities and dissimilarities are prominent among the immediate deliverances of experience. This amounts to the task of showing how the Goodman-Quine account applies to similarity with respect to “sensible qualities”, and, ultimately, how it applies to similarity of sensory experiences themselves. Thus we come back to the central topic of this paper, the notion of “phenomenal similarity”.

The first part of the solution to this problem seems to be provided by the notion of what Quine has called an “innate quality space” and more recently “an innate standard of similarity, or innate spacing of qualities”. Quine remarks that

A response to a red circle, if it is rewarded, will be elicited again by a pink ellipse more readily than by a blue triangle; the red circle resembles the pink ellipse more than the blue triangle. Without some such prior

<sup>13</sup> If what I suggested in footnote 12 is right, genuine similarities consist in the sharing of properties that have the same potential for contributing to the causal powers of the things that have them; and it is certainly at odds with the Humean view to suppose that in being aware of similarities one is *ipso facto* aware of such a causal fact.

spacing of qualities, we could never acquire a habit; all stimuli would be equally alike and equally different. These spacings of qualities, on the part of men and other animals, can be explored and mapped in the laboratory by experiments in conditioning and extinction. Needed as they are for all learning, these distinctive spacings cannot themselves all be learned; some must be innate.<sup>14</sup>

While this point is most clearly illustrated by the case of language learning, in particular the learning of the color vocabulary, Quine stresses that an innate sense or standard of similarity is something we share with animals who lack the capacity for language; it is, as he says, "part of our animal birth-right".<sup>15</sup>

All such learning, according to Quine, involves induction; and the inductions we make depend on the similarities we see, and thus on our innate standard of similarity. The one part of the traditional problem of induction which Quine recognizes as a sensible question is the following:

... why does our innate subjective spacing of qualities accord so well with the functionally relevant groupings in nature as to make our inductions tend to come out right? Why should our subjective spacing of qualities have a special purchase on nature and a lien on the future?<sup>16</sup>

For the answer to this Quine refers us to Darwin, and makes the pithy remark that "Creatures inveterately wrong in their inductions have a pathetic but praiseworthy tendency to die before reproducing their kind".<sup>17</sup>

But now, just what does it mean say that a creature has an innate spacing of qualities that accords with "the func-

<sup>14</sup> "Natural Kinds", p. 123.

<sup>15</sup> *Ibid.*

<sup>16</sup> "Natural Kinds", p. 126.

<sup>17</sup> *Ibid.*

tionally relevant groupings in nature”? Quine would presumably answer this in behavioral terms, referring to how the creature responds, or is capable of responding, to conditioning. And for his purposes such an answer may be entirely satisfactory; I should make clear that in what follows I do not take myself to be raising difficulties for Quine’s account. But for those of us who are not averse to using mentalistic terminology, it is natural to say that what it is for there to be such an “innate spacing” is for the relationships of similarity and difference holding between the person’s perceptual experiences to reflect, or correspond to, relationships of similarity and difference holding between the things experienced. Or, translating this into a somewhat less philosophical jargon, there is such an innate spacing if likenesses and differences in the way things look, feel and sound to the person correspond to similarities and differences between the things themselves. In accordance with the “Goodman-Quine account” of similarity we will suppose that the similarities between the things themselves, the similarities “in nature”, involve the sharing of projectible properties, and that their existence requires that generalizations involving these properties be lawlike. And the fact that we have “immediate” or “direct” perception of these similarities, and the related fact that our inductions based on our experience of things in nature tend to come out right, is to be explained in terms of a sort of preestablished harmony, with natural selection playing the role of Leibniz’s God. (Of course, it is not supposed, as it is by Leibniz, that the experiences are causally unrelated to the things they are of; on the contrary, the existence of the “preestablished harmony” consists in our being causally related to nature in such a way that, *ceteris paribus*, similarities and differences in our experience reflect significant similarities and differences in nature.)

But what of our immediate awareness of the similarities between the perceptual experiences themselves? It is not, indeed, part of the view I have sketched that we must be aware

of these similarities in order to be aware of objective similarities in nature; the view is that perceptual awareness of similarities in nature involves *having* similar experiences, not that it involves *being aware of* similar experiences or of similarities between experiences. Still, by a mere shift of attention one can be aware of similarities between experiences; and as already noted, the awareness of such similarities seems a paradigm of “direct” or “immediate” awareness of similarity. But if we try to explain our awareness of these similarities in the way the account sketched above explains our awareness of similarities in nature (e.g., color similarities), we run into insuperable difficulties.

To begin with, there is the threat of an infinite regress; it seems offhand that we will have to posit experiences of experiences, experiences of experiences of experiences, and so on *ad infinitum*, and that along with this infinite hierarchy of experiences we will have to posit an infinite hierarchy of preestablished harmonies, each consisting in the existence of a general correspondence between the similarity and difference relationships holding between experiences at a certain level and similarity and difference relationships holding between items, namely the experiences at the next level down, of which these experiences are experiences. Clearly this will not do.

There is, indeed, a way of avoiding this infinite regress. Let us suppose that there are “levels” of experiences, the experiences at the first level being experiences of objects in the world, those at the second level being experiences of experiences at the first level, and so on. We can avoid the infinite regress if we are willing to hold that for some level of experiences  $n$  we are able to be aware of similarities and differences between experiences at levels below  $n$  but are not able to be aware of similarities and differences between experiences at level  $n$  itself —for then  $n$  could be the top level, and the regress would stop there. But even then there would be serious, and, I think, insuperable, difficulties. If one is

aware of similarities between experiences by having similar experiences of those experiences, it would have to be possible for there to be a failure of correspondence here, and for one to misperceive the relationships between one's experiences (at any level) in a way analogous to that in which one can misperceive the color relationships between material objects. Just as it can look to one exactly as if two apples are similar in color without their being similar in color, it would have to be possible for it to appear to one exactly as if two color experiences are similar without their being similar — it would have to be possible for it to appear to one exactly as if one were seeming to see two things similar in color without its being the case the one does seem to see two things similar in color. And the latter seems to me clearly not to be possible. I am not claiming that beliefs about one's present visual experiences, about how things presently appear to one or about how one is "appeared to", are "incorrigible". I am merely saying that the ways of being mistaken about one's present experiences, about how things look or sound or feel to one, do not include anything analogous to hallucination or sensory illusion.

Yet it seems clear that we are immediately and noninferentially aware of similarities between perceptual experiences. And we now see that if we take *these* similarities as our "objective" facts, we cannot explain our awareness of them by reference to an innate "spacing" which is subjective relative to them and is, for evolutionary reasons, in accordance with them.

#### IV

This may lead to the thought that similarity of experience is an exception to what was said earlier, following Goodman and Quine, about the connections between the notions of similarity, projectibility and lawlikeness. And this is likely to go with the thought that it is similarity between experiences

that is “intrinsic” similarity *par excellence*. Our first candidate for this status is likely to be color similarity between objects. For the fact that we can directly observe color similarities and dissimilarities in the physical world may at first incline us to think that these are simple and unanalyzable relationships, and that the holding of these relationships is logically independent of considerations of projectibility — since we do not have to *establish* that colors are projectible in order to be aware of these relationships. But on reflection *this* appears as an illusion, fostered by the innateness of our sense of color similarity, that is, by the fact that we are so constituted, presumably as the result of evolution, that in seeing things we have experiences that stand in relationships of similarity and difference that are isomorphic with relationships of color similarity and difference holding between the things seen. Yet it is tempting to suppose that the illusion arises from our projecting onto the color similarity relationships between things the simplicity and unanalyzability of the similarity and difference relationships holding between our perceptual experiences themselves — this would be a manifestation of the propensity of the mind, noted by Hume, to “spread itself on external objects, and to conjoin with them any internal impressions, which they occasion” (*Treatise*, p. 167). The apparent simplicity and unanalyzability of the similarity and difference relationships holding between the experiences themselves cannot in the same way be explained away as an illusion, and it is thus natural to conclude that it is genuine and not merely apparent. And if similarity as a relationship between experiences is unanalyzable, it cannot be analyzed in terms of the notion of projectibility. It should be noted that this way of thinking manifests a not uncommon association of epistemological immediacy and metaphysical simplicity.

But if we reject the Goodman-Quine account of similarity, and reject projectibility as a criterion of similarity and genuineness of kinds or properties, we are left with a problem



which this rejection renders insoluble. Let us say that a property is “grue-like” relative to another property if it is related to it as the “property” of being grue is related to the property of being green. Given any set of properties or features of experiences, we can define properties that are grue-like relative to them, and relative to which they in turn are grue-like. And for any such set of properties we can define a relationship of similarity (or quasi-similarity) that holds between experiences to the extent that they share properties in that set. What is it that determines which of these various properties of experiences are intrinsic, that is, are such that experiences that share them are intrinsically similar, and what is it that determines which of these various relationships of similarity or quasi-similarity is genuine or intrinsic similarity? This question seems to require an answer — and if it can be answered it would seem that the relationship of experience-similarity cannot after all be unanalyzable.

At this point one may be tempted to resort to the idea that the notion of similarity of experience is an undefinable notion which one can come to understand only by being acquainted with cases in which it is instanced. This goes with the idea that this notion, and also the notions of particular features that are intrinsic to experiences, could be introduced, and perhaps could *only* be introduced, by “private ostensive definition”. But I think that it is precisely when we try to divorce the notion of similarity and the notion of an intrinsic feature or kind from the notions of projectibility and law-likeness that it becomes incoherent to suppose that either these notions or the notions of particular intrinsic features or kinds could be defined ostensively. Indeed, it is precisely on the assumption that such a divorce is legitimate that intrinsic features of experience acquire the status of the beetle in Wittgenstein’s box; they become irrelevant, not merely to what we can know about the minds of others, but also to what we can know about our own minds.

It should be noted to begin with that there is a special

difficulty about the idea that the notion of *similarity* might be introduced by a private ostensive definition. For giving an ostensive definition of a concept seems to involve having the thought that the concept applies to the case at hand *and* to cases relevantly *similar* to it, and thus seems to require that the definer already has the notion of similarity. But let us bypass this difficulty for the moment, and consider whether concepts of particular intrinsic features of experiences can be defined ostensively on the assumption that intrinsicness does not logically involve projectibility. I should make it clear that the idea I am attacking is not the idea — which I am sure no one holds — that the intrinsic features of experiences are not projectible; it is the idea that it is at most a contingent fact that the intrinsic features of experience are projectible, and that their intrinsicness has nothing logically to do with their projectibility.

If it is logically an open question whether the intrinsic features of experiences are projectible, it is logically possible that they should not be, and it is logically possible that a given intrinsic feature should not be projectible while one that is grue-like relative to it is projectible. But suppose that I apply a predicate “P” to a given experience, intending thereby to define it as the name of one of the features of that experience, and that subsequently I apply “P” to a number of other experiences. If I do not as yet know whether the intrinsic features of experiences are projectible, can I suppose that I have succeeded in introducing “P” as the name of an intrinsic feature of experiences and that I have correctly applied it to a number of different experiences? I can perhaps suppose that it is possible that I have applied “P” only to experiences having a certain intrinsic feature — but I must allow that if I have done this then either that intrinsic feature is projectible or, if it is not, then it is to some extent an accident or coincidence that I have applied “P” only to experiences that have that feature, in which case it is not *that* feature that I have named (if I have named any). If I think

that I have named some feature "P", I must presumably think that my defining procedure or ceremony, whatever it is, is capable of establishing a connection between an experience of mine having that feature and my thinking that it is right, that is, in accordance with my definition, to apply that term to it. This involves thinking that there are lawlike connections between the instantiation of that feature and other facts. In particular, it involves thinking that something like the following is a lawlike truth: if someone has defined "P" in the way I defined it, then, *ceteris paribus*, if he has any opinion at all about whether "P" applies to an experience of his, he will think that it applies if and only if the experience has the feature in question. Moreover, in order to think that it is possible to establish that a present experience shares a given features with the past experience by reference to which I defined the predicate "P", I must suppose that the feature is one whose past instantiation can be known on the basis of memory or on the basis of inductive inference. But memory provides knowledge of past experiences only on the assumption that there are lawlike connections between certain of those features and the character of subsequent memory impressions —this is the point that the notion of memory is a causal notion. And of course it is only lawlike generalizations that can be directly supported by inductive evidence and used to license inductive inferences concerning the past. All of this indicates that the only features of experiences that I could hope to name by a private ostensive definition are projectible features, that is, those that can enter into the most basic sorts of lawlike connections. Non-projectible predicates can of course be introduced, either by disjunctive definitions like Goodman's definition of "grue" or, in some cases, by exhaustive specifications, by enumeration, of their extensions. But it would seem that the only general features which words can come to name by ostensive procedures are projectible ones. So if I have succeeded in introducing "P" as the name of some feature of experiences,

that feature will be a projectible one. And if I am ignorant as to whether projectible features are intrinsic, I cannot suppose that I have named an intrinsic feature.

Nor can there be any question of my naming a feature ostensively and then finding out empirically that it is intrinsic. To begin with, in order for there to be anything for me to find out I must have some notion of what it is for a feature to be intrinsic, or, what comes to the same thing, what it is for experiences to be intrinsically similar—and of course we are assuming that I am barred from taking projectibility as a criterion of intrinsicness. If the notion of intrinsic similarity is simple and unanalyzable, the only way in which I could introduce it is by a private ostensive definition. But what I have said about ostensive definition in general applies to the special case of the definition of the relational predicate “is intrinsically similar to”. If I take myself to have successfully introduced this predicate by ostensive definition, I must take *it* to be projectible—that is, I must take it as being like the projectible relational predicate “is similar in color to” and not like the nonprojectible relational predicate we would have if we defined the expression “is similar in schmolor to” as being true of pairs of things that share properties that are grue-like relative to colors (properties like Goodman’s *grue* and *bleen*).<sup>18</sup> But in that case I cannot regard it as an open question, and as a matter for empirical investigation, whether intrinsic features—that is, features the sharing of which contributes to intrinsic similarity—are inductively projectible.

It should be remembered that what led to the attempt to divorce the notion of intrinsicness from the notion of projectibility was the fact that we have immediate awareness of what we are inclined to regard as intrinsic similarities between experiences, together with the fact that this awareness

<sup>18</sup> See J. S. Ullian, “More on ‘grue’ and grue”, *The Philosophical Review*, LXX, 3, (July, 1961), 386-389. See also my “On Projecting the Unprojectible”, *Op. cit.*

cannot be explained in the way we can explain our immediate awareness of color similarities between material objects. But this fact, when properly understood, can be seen to be incompatible with the idea to which it naturally gives rise, namely the idea that the notion of experience similarity is "simple and unanalyzable". It will be agreed to be a fact that if someone believes two of his present experiences to be phenomenally similar then, *ceteris paribus*, they will be phenomenally similar. More briefly, though perhaps misleadingly, if experiences seem phenomenally similar, then, *ceteris paribus*, they are phenomenally similar. If this were not a fact, we could not be said to have the ability to be immediately aware of similarities between experiences. But we must ask whether this fact is contingent or conceptually necessary. If it is conceptually necessary, there is a conceptual connection between experiences being similar and their seeming similar to the person who has them, and thus an internal connection between experience similarity and awareness of experience similarity —and if this is so, the notion of experience similarity cannot be a simple and unanalyzable concept. So anyone who holds that this notion *is* simple and unanalyzable will have to hold that there is only a contingent connection between experiences being similar and their seeming similar to the person who has them, as well as holding that there is only a contingent connection between intrinsic similarity of experiences and projectibility. But if someone holds *this*, he is necessarily precluded from having any reason for holding that experiences that seem similar actually are similar, or for holding any other generalization of this sort. He could not hold that he had established such a generalization inductively, for this would require that he already know something which on his own assumptions he could have no way of knowing, namely that the intrinsic features of experiences are projectible and (hence) are such that generalizations involving them are lawlike and supportable by inductive evidence. Nor, for reason already given,

would memory afford him any help. But his situation is worse than this. For his assumptions about the notion of experience similarity (call them Humean assumptions) are such as to preclude the possibility of his *having* a notion satisfying those assumptions. He holds that this notion is not analyzable or definable in terms of other concepts. Nor, on his assumptions, can it be defined ostensively — for we have just seen that it is only on the assumption that the relationship of intrinsic similarity of experience is projectible, and hence that intrinsic features of experiences are projectible, that the notion of this relationship can be introduced by private ostensive definition. So if this notion is assumed to be simple and unanalyzable, and to be such that it is logically an open question whether the relationship of experience similarity is projectible, then there is no way at all in which this notion could be introduced.

Nor — to close a final escape route — would it help to maintain that the notion of experience similarity is innate. If someone claims to have a certain notion, acquired or innate, he must surely suppose that he can know that he is applying that same notion on different occasions. And the Humean assumptions about the notion of intrinsic experience similarity preclude the possibility of anyone's knowing that he has applied this notion on two or more occasions, since, for reasons already given, neither memory nor induction could support such a knowledge claim about an indefinable concept of a relationship that cannot be assumed to be projectible. On these assumptions one cannot claim even to know that certain experiences *seem* similar, let alone that they *are* similar — since in order to know that experiences seem similar one would have to have a notion, that of experience similarity, which on these assumptions no one could have.

The upshot of all this seems to be that projectibility is as much involved in the intrinsic, phenomenal, similarity of experiences as it is in similarities in the physical world, despite the fact that our awareness of the former has a kind

of immediacy that the latter never has, and despite the fact that our awareness of the former cannot be given an explanation that parallels Quine's explanation of our awareness of the latter. But we have yet to give an account of experience similarity that makes this intelligible.<sup>19</sup>

## V

When I mentioned the inverted spectrum problem at the beginning of this paper I remarked that I can not only imagine undergoing spectrum inversion myself but can also imagine behavior that would be evidence that someone else had undergone spectrum inversion. The behavior would include the person's verbal reports, but it would also include what might be called his recognitional and discriminatory behavior; for example, if the person keeps his paper clips in a blue box on his desk and his stamps in an identically shaped yellow box, and if we see him, newly addressed envelope in hand, reaching for the blue box, this will be some evidence that blue things look to him the way yellow things did previously.<sup>20</sup> Those who use the possibility of inverted spectrum to support the claim that sense experiences have phenomenal qualities that are logically independent of their typical causes and effects (and hence cannot be adequately accounted for by causal or functional analyses of mental states) cannot afford to deny that we have this sort of access to the intrinsic, phenomenal similarity of experiences in the intrapersonal case — to do so would be to undercut the only basis there is for thinking that spectrum inversion is possible. But reflection on what is involved in our having this access to the relationship of phenomenal similarity suggests that this rela-

<sup>19</sup> The argument in this section was suggested by some of Wittgenstein's remarks on "private language", but I would not wish to offer it as an interpretation of his so-called private language argument.

<sup>20</sup> For a fuller discussion of how intrasubjective spectrum inversion might show itself in behavior, see my paper "Functionalism and Qualia", forthcoming in *Philosophical Studies*.

tionship itself can be given a causal or functional characterization.

In order to bring this out, I want to consider a fanciful hypothesis, which was suggested to me by Wittgenstein's remark that one should get rid of the idea of a private object by assuming that the object constantly changes, but that one does not notice the change because one's memory always deceives one. According to the hypothesis, the way things look to us with respect to color is constantly changing. At the end of every minute, say, one's spectrum inverts, and a minute later it inverts back again. But this is compensated for by a systematic memory falsification. At any given time one misremembers how things looked during the minute immediately preceding the most recent change, and in every alternate minute before that, in such a way that if a thing has retained the same color throughout a long interval, with the result that the way it has looked to us has been constantly changing, it seems to us, because of the memory falsification, that the way it has looked has been the same throughout.

The question of course is not whether this hypothesis is true or probable but whether it is so much as coherent. And first reactions to this question are likely to be conflicting. One may be inclined to say that it is logically absurd to suppose that someone who can recognize and distinguish things by their colors, and can be counted on to make correct color assignments and correct judgments of color similarity and difference on the basis of what he sees and remembers, is regularly mistaken in such a radical way about the character of his color experience. On the other hand, it seems at first blush that there could be evidence that the hypothesis is actually true of some sub-species of human beings. For example, we might find that people in the sub-species alternate between physiological states  $S_1$  and  $S_2$  at one-minute intervals, and we might have discovered that in our own case changes in these states are correlated with changes in the way things look with respect to color, and we might also find in these



people some periodic change in neural patterns (or whatever) which in our own case we know to be correlated with memory contents. Evidently, then, the question of the coherence of the hypothesis is not one to be settled out of hand.

Since likenesses and differences in the ways things appear to us are supposed to hold in virtue of the intrinsic natures of the experiences we have in being appeared-to in various ways, presumably experiences have intrinsic features corresponding to the observable features of objects, its being in virtue of these features of one's experiences that an object appears to have the features it does. Experiences will not themselves be colored, for we are not thinking of them as sense-data; but we will suppose that there must be some subset of features of experience which stand to one another in relationships of resemblance, complementarity, and so forth, analogous to those relating determinate shades of color, that is, forming a "quality space" isomorphic with the quality space of colors. Suppose, then, that there is such a set of features which is such that there is a constant and systematic change with respect to which of these features a person's experience has when he sees something of a certain color under standard viewing conditions. Let us speak of one color as the *inverse* of another if its position in the inverted spectrum corresponds to the position of the other in the ordinary spectrum, and let us assume that there is a corresponding inverse relationship which holds between the features of experiences on which depend the ways things look with respect to color. Our fanciful hypothesis will now take the form of the supposition that at any given time the seeing of things of different colors goes with the presence of correspondingly different intrinsic features in one's visual experience, but that if at a given time seeing something of a particular color involves one's experience having a given intrinsic feature, a minute later it involves one's experience having the inverse of that feature.

Supposing there to be a set of features of experiences that

alternate in this way, let us call these A-features. What A-features my experience has when I am seeing something blue varies from minute to minute. But using the notion of an A-feature we can introduce the notion of another set of features, which we can call B-features. Letting " $\phi$ " be a variable ranging over A-features, we will say that a person S's visual experiences at times  $t_1$  and  $t_2$  have the same B-features if and only if there is a  $\phi$  such *either* (1)  $t_1$  and  $t_2$  are within one-minute intervals which both began either an even number or an odd number of minutes before or after a certain time T (say Noon today), and S's experience had  $\phi$  at both  $t_1$  and  $t_2$ , or (2)  $t_1$  and  $t_2$  are within one-minute intervals one of which began an even number of minutes before or after T and the other of which began an odd number of minutes before or after A, and S's experience at  $t_1$  had  $\phi$  and his experience at  $t_2$  had the inverse of  $\phi$ . A-features and B-features are thus so related, by definition, that the B-features of someone's visual experiences over a period of time remain the same just in case the A-features are replaced by their inverses at one-minute intervals, and *vice versa*.

But now let us consider again the idea that it might be true of some sub-species of human beings, or perhaps even of us, that the way things look to members of that species with respect color is constantly changing, and that this is compensated for, and goes unnoticed because of, a systematic memory falsification. We have supposed, or rather stipulated, that there is a constant and systematic change with respect to what A-features a person's experience has when he sees things of certain colors, and we can allow that this amounts to there being a sense in which the way things of certain colors look to us is constantly changing. But given that this is so, what possible reason could there be for saying that it is in *this* sense that we use "appear" and related expressions in our memory statements? Let us mark with the subscript "A" the sense of "appears" and related expressions in which sameness in the way things appear, and sameness of ways of being

appeared-to, goes with sameness with respect to A-features, and let us mark with the subscript "B" another sense of these expressions in which sameness of ways of being appeared-to goes with sameness with respect to B-features. Given our supposition (or stipulation) about A-features, and what it implies about B-features (given my definition), cases in which I *think* that I am appeared-to in the same way as on a past occasion are normally cases in which I am appeared-to<sub>B</sub> in the same way as on that past occasion. Moreover, it is precisely when I am in the same state of being appeared-to<sub>B</sub> on different occasions that I am able to recognize things by their similarities with respect to color, and it is precisely when things appear<sub>B</sub> different to me that I am able to distinguish them by their differences with respect to color. Given this, there seems every reason to say, and no reason not to say, that it is sameness and differences of states of being appeared-to<sub>B</sub>, not sameness and difference of states of being appeared-t<sub>A</sub>, that corresponds to "intrinsic" phenomenal likeness and difference of experiences. And this makes it seem totally perverse to suggest that my memories of how things of various colors appeared to me in the past are memories of how they appeared<sub>A</sub>, and thus false as often as true, rather than memories of how they appeared<sub>B</sub>, and so mostly true; the perversity is comparable with that of taking and ordinary person to mean by "green" what Goodman means by "grue", and judging his inductive inferences accordingly. We can tie this up with the discussion in the preceding section by saying that if someone introduced by "private ostensive definition" a name for one of the "intrinsic features" of his visual experiences, there could be no reason for saying that he had introduced it as the name of an A-feature.

Given a sense of "appear" in which it is true that the way things of certain colors appear is constantly changing, or is different in one part of the visual field rather than another, we can introduce, by means of the sort of definition em-

ployed above, a sense in which this is not true; and given a sense in which it is not true, we can introduce any number of senses in which it is true. So it is really a truism that there are possible senses in which it is true and a possible senses in which it is not. The important question is what determines which of these possible senses is the actual sense, the sense in which we employ the words "appear" and "look" when we say that in fact the way blue things standardly appear remains the same and that only if there occurred phenomena of a very special sort—a certain sort of temporary breakdown in recognitional and discriminatory capacities, together with appropriate verbal reports—would it be reasonable to say that there had been a change in the way blue things standardly look to someone? And my suggestion is that what singles out one of these possible senses as the actual one is its being such that it is likenesses and differences in the way things appear to someone in that sense that explains his recognitional and discriminatory abilities.

One way of putting this view is by saying that the notion of similarity of experience must be understood in terms of the more fundamental notion of experience *of* similarity, and, likewise, that the notion of dissimilarity of experience must be understood in terms of the notion of experience of dissimilarity. Thus we might say, as a first approximation, that experiences are similar if they jointly yield awareness of similarity, or, better, they that are similar if they would yield awareness of similarity if they were "co-conscious", that is, conscious to a person at the same time, where an experience counts as conscious to a person when he correctly remembers it as well as when he is actually having it. This account would of course be circular if the only similarities of which awareness is yielded were similarities between the experiences themselves—we would then be saying that what it means to say that experiences are similar is that they yield or tend to yield awareness of the similarity of those very experiences. The account will not be circular,

however, if the similarity referred to in the *analysans* is, or includes, similarity of something other than experiences. So it is a possible view, and I think basically a correct one, that what it means to say that experiences are phenomenally similar is that they stand in a relation which is such that if two experiences are related by that relation, and are co-conscious, their joint occurrence yields, or tends to yield, awareness of similarities holding between material things, namely similarities with respect to their perceptible or observable qualities, the latter application of the notion of similarity, namely to material things, being explicable in terms of the Goodman-Quine account as involving projectibility and lawlikeness.

Phrases like "experience of similarity" must here be understood intentionally. For there to be similarity of experience, and so experience of similarity, there need not be an objective similarity in the world that is veridically perceived; it is enough if it looks as if, or sounds as if, or feels as if there were such an objective similarity, or that it would look, sound or feel this way if the experiences were co-conscious. On this view the connection between similarity of experience and projectibility and lawlikeness is rather complex. On the one hand it can be seen as an intentional connection; similar experiences are, when co-conscious, jointly *of* objective similarities which may or may not actually exist in the world but whose actual existence would involve the sharing of projectible properties. But in virtue of this connection; similar experiences are, when co-conscious, jointly in virtue of which they hold, are themselves projectible. Indeed, something stronger than projectibility is involved here. To say that features and relationships are projectible is not to say that they actually are connected in a lawlike-way with other features; generalizations involving such features and relationships are lawlike in the sense that they would be confirmed by positive instances, but this is not to say that any such generalization is actually confirmed to any extent,

or that any such generalization is actually true. But given the proposed account of similarity of experience, it appears that there must actually be lawlike connections between experience similarity and other states of affairs. Specifically, there must be a lawlike connection between a person's having phenomenally similar experiences and his believing that there exist, and that he is perceiving, objective similarities in the world, and likewise, there must be a lawlike connection between a person's having phenomenally different experiences and his believing, or tending to believe, that there exist, and that he is perceiving, objective differences in the world. This does not mean that, for example, whenever similar experiences are co-conscious there results a belief in objective similarity; rather, we can say that co-conscious experiences that are phenomenally similar *tend* to bring about such beliefs, and do so in the absence of "countervailing" factors.<sup>21</sup>

## VI

But where does this leave our awareness of similarity and difference of experience? How does defining phenomenal similarity and difference of experience in terms of their causal role in awareness of objective similarities and dif-

<sup>21</sup> When I read this paper in Mexico City, Mr. Gareth Evans pointed out in the discussion that there is an apparent circularity in this account. Since any two objects will be similar in some respect or other, and since any reflective person knows and believes this, we can hardly characterize the belief to which similar experiences give rise as simply the belief that the perceived objects are similar in some respect or other. But (so the objection runs) any characterization of the belief which is more specific is bound to bring in reference to perception, and, ultimately, to similarity of perceptual experiences. Thus we might characterize the belief as the belief that two objects are similar with respect to their perceptible properties. But what is a perceptible property, if not one whose instantiations tend to give rise to similar experiences in perceivers? Or one could perhaps characterize the belief as the belief that there is a similarity between the objects which accounts for the relation between the experiences of them — and what is the relevant relation between the experiences, if not their similarity? (This formulation of the objection is mine, and may distort what Evans had in mind.)

I agree that the characterization of the belief is bound to bring in, eventually, reference to perception and perceptual experiences — indeed, this is closely connected with the essential "reflexivity" of perceptual awareness which I briefly discuss in section VI. But it is not clear to me that it will

ferences explain our awareness of phenomenal similarities and differences themselves? To this question I have no fully satisfactory answer, and certainly none that I can give briefly. What I wish to contend is, to speak both vaguely and metaphorically, that our ability to be aware of experience similarities is implicit in, and is a sort of shadow or reflection of, our ability to be perceptually aware of objective similarities in nature. A crucial point here is that while there are ways in which a man or animal may fail to be aware of experiences he has, none of these is analogous to blindness. If, as I suppose, we are disposed to deny that dogs are aware of their experiences, this is not because we think that dogs lack an inner sense that we possess, but because we think that they lack concepts that such awareness would require. It makes no sense, I think, to suppose that there might be a creature of human intelligence who has these concepts, and who has perceptual experiences by means of which he perceives objective states of affairs, but who differs from us in lacking the ability to be aware of these experiences and

have to bring in reference to phenomenal similarity of experiences. Suppose we characterize the belief as the belief that there is a similarity between the objects which accounts for the relation between the experiences of them. While this relation will be, or include, their phenomenal similarity, it is not clear why this fact must be mentioned in the analysis. Again, while it may be necessarily true that instantiations of perceptible properties tend to give rise to similar experiences in appropriately located observers, it does not obviously follow that we cannot explain the notion of a perceptible property without explicitly invoking the notion of phenomenal similarity of experience.

But I would not be surprised if it could be shown that the circularity is unavoidable. For I think it is true in general that philosophical "analyses" of central concepts (causality, substance, property, etc.) are bound to be circular in this way, and that such circularity is not necessarily a defect. I cannot say in general what distinguishes "vicious" from "virtuous" circularity; but I think that I can sometimes tell the difference in particular cases. (There is an obvious difference between my account and the blatantly circular account which says that experiences are similar if they tend to give rise to the belief that they are similar. In both cases (let us suppose) the notion being analysed is invoked in the *analysans*. But in my account various other concepts are invoked as well, whereas in the other the only other concept that is invoked is the concept of belief — and there is something peculiarly unacceptable about any account that purports to explicate the notion of P's being the case in terms of the notion of someone's believing P to be the case.

of relations of similarity and difference between them. But this, I think, is to be explained in terms of what is involved in perceptual awareness of objective states of affairs. The propositional content of a perceptual awareness always involves a token reflexive element and, implicitly, a reference to the perceiver and his experiences. How this content should be articulated depends on the conceptual sophistication of the creature doing the perceiving. At relatively primitive level the content might be expressed in some such words as "Tree here now". But for a creature who has the concept of himself, and the concept of seeing, a more adequate expression would be "I see a tree in front of me". This might be put by saying that in perceiving something to be the case one is (given human intelligence and conceptual ability) at least potentially aware that one perceives it to be the case. Here again no analogue to blindness is possible. It is not by an inner sense, which I might have lacked without lacking the ability to see, that I am aware that I see. It would scarcely be intelligible for someone to say "I know that there is a tree in front of me, but I have no idea whether I see a tree in front of me". But if seeing involves having visual experiences, then the ability to be aware that one sees, which is (I have claimed) involved in the ability to see, involves the ability to be aware of one's visual experiences. And I would hope to be able to show that the ability to be aware, in the required sense, of perceptual experiences involves the ability to be aware of phenomenal similarities and differences between these experiences. But that, and the elaboration of the account crudely sketched in these last few pages, is a matter for another paper.<sup>22</sup>

<sup>22</sup> This paper is based on material I presented as part of my John Locke Lectures at Oxford University in the Spring of 1972. The present version was written during 1974 at the Center for Advanced Study in the Behavioral Sciences (to which I hereby express gratitude) and was first read at the Western Washington Philosophy Colloquium at Bellingham, Washington, in March, 1974. It has also been read at the Davis and Santa Cruz campuses of the University of California, and at the University of Mexico. I am grateful to Jonathan Bennett for very helpful comments on an earlier version.



## RESUMEN

Shoemaker comienza citando un punto de vista defendido por D. Locke y por A. Donagan, según el cual los conceptos mentales tienen un doble juego: por una parte refieren a algo privado, por otra, refieren a algo público. Así, por ejemplo, 'Dolor' es a la vez una sensación y la conducta de quejarse, etc. Esta teoría parece hacerle justicia al carácter inmediato de lo mental (algo privado) y al uso público de los conceptos mentales. De otra parte la teoría parece acomodar casos como el del espectro invertido en el cual una persona puede ver lo rojo como verde y así con los otros colores. Sin embargo, la teoría tiene un resultado implausible, a saber, acepta la posibilidad de que alguien sienta como dolor lo que expresa como una cosquilla.

Esta teoría parece requerir similitudes fenomenológicas inmediatamente dadas que están relacionadas sólo contingentemente con la conducta o los estados corporales. Por lo tanto, si esta idea de la similitud se demuestra viciosa se habrá refutado la tesis del doble juego de los conceptos mentales.

En la segunda parte Shoemaker expone su interpretación de la tesis sobre similitud que comparten Goodman y Quine. Ambos filósofos rechazan toda teoría innata o subjetivista de la similitud y dan un análisis objetivo de la misma en términos de la inducción. Ahora bien, hay inducciones con términos como 'verde' que son válidas y otras con otros términos que no lo son. ¿Qué nos permite decidir cuáles son válidas? Goodman dice que el criterio de validez lo da la proyectabilidad de un término v.g. 'verde', las generalizaciones en las que aparece un término proyectible son válidas para casos no examinados aún. Luego, los objetos que comparten un término proyectible forman clases naturales. Así, pues, la proyectabilidad nos da la similitud de un predicado o propiedad y ésta alcanza así, una medida objetiva: nuestras inducciones dependen de esas proyecciones y a su vez nuestra supervivencia depende de esas inducciones. Una vez que se da este análisis de la similitud en términos de la proyectabilidad de predicados, podemos deshacernos del indeseable concepto de similitud.

En la tercera parte examina Shoemaker un aparente contraejemplo a la tesis antes expuesta: el caso de la similitud entre experiencias. Esta similitud está entre los datos inmediatos de la conciencia y parece que no se puede reducir a la proyectabilidad de pre-

dicados como en el caso de los objetos materiales. En el caso de éstos lo que permite el análisis de similaridad en términos de proyectabilidad es el hecho contingente de que los seres humanos tenemos un espacio innato de cualidades que permite su discriminación y a partir de él hacemos las inducciones; este espacio innato concuerda (esto también es un hecho contingente) con las agrupaciones funcionales que hay en la naturaleza y este acuerdo explica, en última instancia, nuestra supervivencia. Ahora bien, la cuestión surge: ¿cómo explicar el acuerdo entre la experiencia inmediata de similaridad y las experiencias perceptuales mismas, por ejemplo? El dilema aparece: o bien, enfrentamos un regreso infinito con una infinita jerarquía de experiencias o aceptamos la consecuencia de que hay experiencias que parecen similares sin serlo.

La razón de esta insatisfactoria conclusión es que en la similaridad entre experiencias a diferencia de la similaridad entre objetos materiales no hay el contraste entre lo subjetivo y lo objetivo: la similaridad es subjetiva y las experiencias también lo son y por ello no se puede excluir la total arbitrariedad. Dicho en otras palabras: no se puede hablar de concordancias entre un espacio innato de cualidades y las agrupaciones de objetos porque en el caso de las experiencias no hay esas dos instancias. ¿Debemos aceptar entonces que la similaridad entre experiencias constituye un contra-ejemplo al análisis presentado en la segunda parte? Un consuelo sería decir con Quine que después de todo es sólo un hecho contingente que tengamos experiencia de similaridades. A Shoemaker no le satisface esto y aunque acepta que ése es un hecho contingente dice también que en tanto tengamos experiencia de similaridad debemos dar una explicación de la misma y esa explicación debe ser en términos del análisis expuesto en la segunda parte. A esto se dedica Shoemaker en la cuarta parte.

Shoemaker parte del caso de la similaridad entre colores, la cual parece simple e inmediata y por tanto parece no envolver proyectabilidad. Pero esto no es así, pues entre dos similaridades opuestas debe haber una decisión acerca de cuál de las dos es genuina. Una explicación se ofrece: es genuina aquella que se introduce mediante una Definición Ostensiva Privada. Pero esta explicación enfrenta la siguiente dificultad: para determinar cuál similaridad es genuina debe haber algo más que la experiencia, pues de otra forma no se podrá introducir esa similaridad en razón de que no habrá más que la experiencia introducida por esa definición y la experiencia en cuestión no lleva en sí misma la similaridad  $X$  o  $Y$ . La Definición Ostensiva Privada no elimina, entonces, la arbitrariedad.

Shoemaker se dedica luego a mostrar que, en general, no se pue-

de introducir la similaridad entre experiencias sin introducir también la proyectabilidad. Así desecha los intentos de introducir primero la similaridad para después encontrar que la propiedad es proyectible; también desecha el recurso a la memoria, pues la causalidad de ésta resulta incompatible con el divorcio entre similaridad y proyectibilidad y finalmente rechaza la posibilidad de apelar a la inducción. Por último, presenta un argumento según el cual si se supone la contingencia de la conexión entre los conceptos de similaridad y proyectibilidad no se podrá llegar a la generalización.

‘Las experiencias que parecen similares son similares’ y no la inducción ni la memoria ni las ideas inatas pueden prestar ayuda siempre. Por lo tanto, debe rechazarse la tesis de que la similaridad entre experiencias es algo simple e inmediato. Sin embargo, esto no prueba que la similaridad entre experiencias deba explicarse de otra forma. Shoemaker presenta entonces —en la quinta parte— dos argumentos en favor de la tesis de que la similaridad entre experiencias debe analizarse en términos de las similaridades objetivas que son los correlatos intencionales de aquellas.

En el primero, Shoemaker examina un caso hipotético según el cual las notas de las experiencias sensoriales variarían a intervalos regulares y se pregunta: si esto sucediera y hubiese varias posibilidades de variación ¿cuál de ellas sería la que funcionaría en los juicios de percepción? Es necesario introducir un criterio para decidir esto y el criterio es éste: las notas que funcionarían serían aquellas que explican nuestras capacidades reconocionales y discriminatorias. La decisión entonces, respecto de la similaridad de experiencias la determina el funcionamiento de esas capacidades y por lo tanto, el mundo objetivo.

El segundo argumento es más fuerte y establece que a menos que las similaridades entre experiencias se analicen en términos de similaridades objetivas no tendremos un análisis de esas similaridades, pues siempre caeremos en explicaciones circulares. Pero, ¿cómo dar el análisis de la experiencia de similaridad en términos de similaridades objetivas? Shoemaker pospone la respuesta y advierte en la sexta parte, que al presente lo que le importa es rechazar una cierta imagen de esas similaridades según la cual carecer de la similaridad entre experiencias es equivalente a estar ciegos o a no distinguir algo ante los ojos debido a la penumbra. Por el contrario, establece Shoemaker, carecer de esas similaridades sería equivalente a perder esas capacidades en las que nos va la vida lo cual es ciertamente algo sustancial.

Resumen por  
Enrique Villanueva