

## 1

## An Approach to Explanation

*The word explanation occurs so continuously and has so important a place in philosophy, that a little time spent in fixing the meaning of it will be profitably employed.*

J. S. Mill

## 1

Philosophers of science should be able to state the truth-conditions that govern the occurrences of 'to explain' when that verb functions as in, for example,

- (1) The kinetic theory of gases explains why the state equation of vapors near condensation differs markedly from that of an ideal gas.

These truth-conditions determine what must be true of any doctrine or proposition rightly said to explain something. Thus they constitute standards that play a crucial part in the critical analysis of any alleged piece of information or theory, and their explicit formulation is needed by anyone who wishes to know in general when and how specific scientific contributions meet these standards.

What are these truth-conditions? Most answers proposed up to now strike me as missing the heart of the matter, or as being at best about quasi-technical homonyms of 'to explain'. In this essay I shall propose an answer of my own. It too has weaknesses: it is unfinished and it raises a number of problems whose solution is not available at present. I think, however, that it sets matters in the right perspective and that it provides foundations for further inquiries.

'To explain' does not always function as in (1). Consider the following two statements:

- (2) Someone explained to somebody how World War II might have been avoided.
- (3) Newton explained a long time ago the variations of the tides.

Statement (2), unlike (1), can generate the questions 'where?', 'when?', 'using what words?', that is, (2), unlike (1), implies the occurrence of some datable and placeable exchange of utterances. (3), on the other hand, can give rise to 'when?' but not to 'using what words?'

The truth-conditions of statements in which 'to explain' functions as in (2) or as in (3) must obviously be different from those in which it functions as in (1). They are also of less direct concern to philosophers of science. In what follows I shall nevertheless give them a good deal of attention. I shall, in fact, begin with a long discussion of statements like (2). Nothing very deep lies behind this. It happens to be a convenient way of introducing certain notions that are at the heart of my approach.

Toward the end of this essay I shall turn to 'explanation', and explain what an explanation is.

A word of caution. The verb 'to explain' and the phrase 'explanation of' are often followed by the oratio obliqua of certain questions and these questions fall naturally into groups: one can explain why, or how, and have explanations of what really happens whenever such and such, or of what causes what, and so on. Many writings ostensibly about explanation are best understood as not about explanation at all but as about what is called for by the questions belonging to one or the other of these groups. That should not be true of this paper. It is *not* about causality, or reality, or why, or how; it is about explaining and explanation and only about that.

## 2

Let us begin by asking what it is for someone to explain something to someone, where the "something" can be specified by means of an indirect question.<sup>1</sup> In other words, what are the truth-conditions common to such statements as 'John is explaining to Rosalie how

<sup>1</sup>"INDIRECT QUESTION is the grammarian's name for a modification of what was originally a question, such that it does not stand by itself as a sentence, but is treated as a noun serving, for instance, as subject or object to a verb outside of it. Thus the direct question *who are you?* indirect question, *I asked who he*

ozone differs from ordinary oxygen', 'Rudolph explained to Susan what "functor" means', 'He will explain to you what you should do'? More formally, what are the truth-conditions governing statements of the form 'A E to B W', where A and B indicate positions occupied by expressions through which a person or persons are mentioned (that is, proper names, singular descriptions of specific people or groups of people, personal pronouns, 'somebody', 'someone', etc.) where W indicates a position occupied by an indirect question, and E indicates a position taken up by a tensed form of 'to explain'.

Two things may be noted immediately: (a) Many kinds of indirect questions can occupy the position indicated by W, and they may open with a variety of interrogatives—'Why', 'How', 'Whence', 'Whither', 'What'—but not *every* indirect question is at home there; some would be out of place, awkward, reminiscent of Eisenhower prose, for example, 'what the distance in miles between London and Paris is' or 'whether it will rain tomorrow' or 'what time it is' or 'which sister Sam married'. A good analysis should show why some indirect questions do not sit well in these contexts.

(b) In these statements, 'to explain' functions as an accomplishment term. 'Accomplishment term' is used here in the sense given to it by Vendler (1957).<sup>2</sup> It indicates a place within a fourfold classification: activity terms, accomplishment terms, state terms, achievement terms. What follows is a brief account of this classification.

was ..." (Fowler 1954, p. 268). I shall use 'indirect question' and 'question in oratio obliqua form' interchangeably.

<sup>2</sup>I believe that I have reproduced the distinctions that Vendler must have had in mind when he wrote his paper, but I do not rely on his criteria. He writes "... if I say of a person that he is running a mile or of someone else that he is drawing a circle, then I do claim that the first one will keep running till he has covered the mile and that the second will keep drawing till he has drawn the circle. If they do not complete their activities, my statement will turn out to be false." This is a mistake: if either the runner or the drawer should be interrupted before finishing, the statement will not "turn out" to be false. I have tried to avoid such objections by attending to the use of the simple past. But Vendler saw what had to be seen.

Note that the rest of the discussion does not require that the classification be exhaustive. Nor does it require that a given verb or verbal phrase occur as a member of the same class in every context. 'To explain', as we shall see, does not *always* function as an accomplishment term. The principle at work here is analogous to that in virtue of which some verbs can be transitive in some types of contexts but intransitive in others.

The first two classes, activity terms and accomplishment terms, include all those verbs that have a continuous (progressive) present or, more exactly, those whose continuous present makes sense as the main verb in answer to 'what is *x* doing?', for example, 'to run', 'to eat', 'to talk', 'to lunch', 'to speak'. They are used to describe episodes that require and occupy stretches of time, stretches at any moment of which the continuous present applies. The last two classes exclude these.

The difference between activity terms and accomplishment terms is readily seen when we compare their simple past tense (perfective). Both have a simple past tense that implies that the continuous present was applicable at some moments in the past. The simple past tense of an activity term is applicable as soon as such moments have passed, and implies only the existence of such moments in the past. Aristotle walked. This implies that at some moments in the past Aristotle was walking. It does not tell us whether or not Aristotle is through walking. The simple past tense of accomplishment terms implies more. It implies that relevant activities took place in the past, but furthermore that they have come to an end. And not to a mere stop, but to a conclusion. In other words, the simple past tense of accomplishment verbs entails that something has been finished, completed, something that might, in principle, have been left unfinished, incomplete, that might have been stopped before coming to its logical end. Typical accomplishment verbs are 'to fill', 'to draw a circle', 'to write a letter', 'to die'. During the appropriate sorts of goings on, it is true that Jane is filling a pillow, Zachary is drawing a circle, Rachel is writing a letter, Socrates is dying. But to say that Jane filled the pillow, that Zachary drew a circle, that Rachel wrote a letter, that Socrates died, is to imply that all the stuffing is in, that all the points of the perimeter have been joined, that the final salutation is on paper, that he is dead. Otherwise, all that may be said is that she was filling a pillow, that he was drawing a circle, that she was writing a letter, that he was dying.

Thus accomplishment terms differ from activity terms in being associated with distinctions between what constitutes completion and what constitutes mere stopping, mere interruption. As Vendler points out, activity terms lend themselves to such questions as 'How long did so and so do such and such?', for example, 'How long did Aristotle walk?', but accomplishment terms call instead for 'How

long did it take so and so to do such and such?', for example, 'How long did it take Jane to fill the pillow?'<sup>3</sup>

The difference between state terms and achievement terms is of less immediate relevance. Briefly, state terms (for example, 'to know', 'to love', 'to want', in their nonepisodic senses) have present indicatives that apply throughout intervals of time, and that are true of something at more than one moment even if true of it only once; achievement terms (for example, 'to win', 'to hit', 'to join', in their nonepisodic sense) have present indicatives that are true of something at one particular moment only, unless they are true of that thing more than once.

It should be clear that in 'A E to B W' contexts, 'to explain' functions as an accomplishment term: it can appear there in the continuous present; episodes filling intervals during which a continuous present applies need not but may be followed by moments at which truth demands the substitution of the simple past; finally, the use of the simple past does imply that the agent mentioned as subject of the sentence completed something.

A good analysis should therefore make explicit the nature of the completion implied by statements in 'A E to B W' form in which 'to explain' occurs in the *simple past tense*; it should bring out what must be the case for a statement in 'A explained to B W' form to be true.

Let us now turn to such an analysis. To expedite the discussion, I shall refer to episodes whose completion can be conveyed by a statement of the form 'A explained to B W' as *explaining episodes*. I shall also refer in a general way to whatever person or persons would be mentioned at A as the *tutor*, and to whatever person or persons would be mentioned at B as the *tutee* in the episode; when I speak of 'the question', I shall mean the question whose oratio obliqua form would appear at W as the complement of 'explained'.

In what follows I shall examine—and reject—a series of hypotheses. Each offers a tentative and tempting analysis. My purpose in following this procedure is twofold: to bring out negative truths that are of intrinsic interest, and to discover the conditions that an ac-

<sup>3</sup>Verbs that in isolation function as activity terms may, of course, be constituents of verbal phrases that function as accomplishment terms, for instance, 'to write', as in 'to write a letter', 'to walk' as in 'to walk to the store'.

ceptable doctrine must satisfy. Each demonstrable defect of these tentative analyses will correspond to such a negative truth and to such a condition. Confidence in the last hypothesis will have to rest to a great extent on the fact that it has none of the defects of its predecessors and that it apparently has none of its own. Some fatal flaw may of course have escaped my attention. But I can think of no procedure that will yield conclusive results here. This one has a certain advantage: If defects of the last hypothesis do eventually come to light, they will show that I stopped too soon, but they will not vitiate the negative conclusions nor invalidate the negative conditions already established.

*First hypothesis:* A statement of the form 'A explained to B W' is true if and only if an episode has occurred

- (a) at the beginning of which the tutee did not know the right answer to the question;
- (b) in the course of which the tutor presents to the tutee all the facts that he, the tutee, must know to know the right answer to the question;
- (c) at the end of which the tutor has presented all the facts mentioned in (b);
- (d) at the end of which the tutee knows the right answer to the question.

The hypothesis avoids the implication that in every explaining episode someone actually asks the question. This is as it should be. The truth of 'Salviati explained to Sagredo what the grounds are for saying that similar solids are to each other in the sesquilateral ratio of their surfaces' does not imply that Sagredo asked the question, though we can discover by reading the dialogue that he did.

The hypothesis nevertheless still contains unnecessary conditions. Initial ignorance on the part of the tutee is not necessary, is not implied: we all know of the father who explained to his son how babies come into the world only to be told "I knew all that; I thought that you were going to tell me something about birds and bees!" Nor is ultimate knowledge on the part of the tutee necessary: "Max explained to Alonzo what the difference is between using and mentioning a term" does not entail that Alonzo now knew, that he was convinced, or that he even paid any attention.

Eliminating (a) and (d) and rephrasing (b) and (c) as

- (b') in the course of the episode, the tutor presents the facts that one must know to know the right answer to the question;
- (c') at the end of the episode, the tutor has presented all the facts mentioned in (b');

would avoid these unwanted implications, but it would not take care of another objection—that the hypothesis (or its revised version) does not provide a set of *sufficient* conditions. At twelve o'clock yesterday, I did not know the height of the Empire State Building; I met Henrietta, who knew and who told me "1,250 feet." Now I also knew. This episode meets all the requirements of the first hypothesis, but it does not rate the description 'Henrietta explained to me what the height of the Empire State Building is'.

Of course, no episode can rate the description, 'Henrietta explained to me what the height of the Empire State Building is'. This may suggest that the sentence is perhaps not well formed, that it is grammatically rather than semantically deviant. The point is not without interest since we want to know why some questions do not function properly as direct objects of 'to explain'. However the following two illustrations bring out that more than syntax is at play here; they show that the same string of words is a proper direct object of 'to explain' in some circumstances but not in others.

First illustration: maximum security prison, walls that can't be scaled, guards in every hall, dogs, and so on. A prisoner manages to escape by digging a tunnel. The tunnel has been found, but no one can figure out how the dirt from the excavations was disposed of; all the possibilities that occur to anyone have to be rejected for one reason or another. The prisoner is now recaptured. He is made to describe his escape. He is clearly aware of the fact that his jailers are baffled, and in the course of his description, he reveals how he disposed of the dirt from his tunneling.

I take it that there can be no objection to saying that the prisoner explained to his jailers (in the course of his account) how he disposed of the dirt from his tunneling.

Second illustration: similar jail, similar escape, but in this case it is easy to see how the prisoner may have disposed of the dirt; he may have dropped it into the moat under his window, or he may also have carried it out in his pocket to a garden where he was not

closely watched. But still, the jailers do not *know* how the prisoner disposed of the dirt, whether he used the moat or his pockets. The prisoner is recaptured, made to describe his escape, and he reveals how he disposed of the dirt (the moat) from his tunneling.

Here, as in the previous example, the conditions of the first hypothesis are satisfied, but I take it that 'The prisoner explained to his jailers how he disposed of the dirt from his tunneling' is out of place, distorts things, smacks of exaggeration, is at best a near truth.

It would not do, of course, to say simply that the prisoner did not explain to his jailers how he disposed of the dirt from his tunnel, and to let it go at that. This would suggest that he refused to explain, or failed to explain, or prevaricated, or perhaps did not address himself to the question at all. Normally some such thing is conveyed by the explicit denial that someone explained something on some specific occasion. Thus in the present instance that sort of denial would be misleading. The description does not apply, but neither does its denial. Not an unusual situation; one does not want to say simply that the Empire State Building is not happy, for this would suggest that it is melancholy. But one can't say that it is happy either.

Admittedly, the statement that in the second illustration the prisoner explained how he disposed of the dirt from his tunneling ceases to be glaringly odd when one limits one's mental picturing to the scene in which the prisoner describes his escape; more of the epistemic background must be kept in mind, that is, that what was said, though informative, merely settled for the audience *which* of the means of which it was already aware had actually been used, and that the prisoner realized this.

It should not come as a surprise that the first hypothesis is inadequate. A correct perspective on the nature of explaining episodes would reveal the difference between what is conveyed by statements of the form 'A explained to B W' and statements of the form 'A told B W'. The first hypothesis would obliterate that difference.

The first hypothesis requires that what the tutor tells the tutee be true. However, there is such a thing as explaining incorrectly, and a report that little Suzy explained to her brother what little girls are made of does not imply that what she said is true. Seen in this light, the requirement may seem uncalled for, but this would be an illusion. Statements of the form 'A explained to B W' *do* imply that

the tutee was given correct information. The addition of 'incorrectly' does not add a detail that could have been left unspecified without affecting the truth of the statement; it *qualifies* the whole statement, and has the force of a partial denial: to have explained incorrectly is *not* to have explained, but to have been engaged in something akin to it. The same effect can be achieved through ironic, or sarcastic, or condescending, intonation.

Admittedly, statements of the form 'A explained to B W' are open to strict and to loose usage. Strictly, they imply that the tutor uttered verities. But one does not have to know that he did, to use such a statement; it is enough to assume that he did, and a loose usage even allows the tutor's views about the truth of what he said to prevail over those of the user. It does not rule out such statements even when the latter believes that what the tutor told the tutee was false. Such usage, however, is apt to mislead when one's listener does not know better.

The hypothesis also entails that the question itself must be sound, that is, must have a correct answer. Here too there is a strict and a loose usage. The former rules out questions that are based on false presuppositions, the latter countenances them. 'Max explained to Alonzo why the Gödel Incompleteness Theorem is of no epistemological significance' would normally commit the user of the sentence to the view that the Gödel Theorem is of no epistemological significance: the statement, unless qualified, implies this. Implications of this kind are stillborn whenever speaker and listener agree, and know that they agree, that the question is unsound.<sup>4</sup>

### P-predicament

The statement of the second hypothesis must be postponed until the term 'p-predicament',<sup>5</sup> needed for its formulation, has been introduced.

Consider the following two questions:

<sup>4</sup>The discussion from here on will be limited to the strict usage. The looser usage is best handled after the analysis of the strict usage, but I will ignore it altogether as of little interest.

<sup>5</sup>The p of 'p-predicament' may be thought of as standing for 'puzzled' or 'perplexed', but mnemonically only. As will become obvious, one need not be either puzzled or perplexed to be in a p-predicament. (In more current terms, statements of the form 'A explained to B W' entail that the tutor held W to be sound and his answer to be true; they implicate, in Grice's sense, that the user agrees.)

- (4) What is the height of Mt. Kilimanjaro?
- (5) Why do tea kettles emit a humming noise just before the water begins to boil?

I know the answer to neither. However I am not equally ignorant with regard to both. This is brought out to some extent by the fact that I am prepared to say that *I do not understand* why tea kettles emit a humming noise just before the water begins to boil, whereas I am not prepared to say anything of the sort about what the height of Mt. Kilimanjaro is. We will say that I am in a p-predicament with regard to (5) but not with regard to (4).

What is the difference?

I take (4) to be a sound question, that is, to admit of a right answer,<sup>6</sup> to rest on no mistaken assumptions or false presuppositions; there is such a thing as Mt. Kilimanjaro, and it is the sort of thing that has a height. I also believe the following about the answer: (a) it can be put as a number followed by a unit of length; (b) if given in feet, the number must lie between 100 and 30,000; (c) if not in feet, then the number must be such that conversion into an equivalent number of feet would yield a number between 100 and 30,000; and (d) the number could be obtained through certain measuring operations that need not be described here. These conditions on the answer, even when combined with my other beliefs, do not uniquely determine an answer to (4); acceptance of them does not settle the question for me. But acceptance of them does require that I consider a number of expressions as not possibly being formulations of the right answer. These make up a varied lot: '12 feet', 'I don't know', 'Look it up in the encyclopedia', 'Morton White', '19,321 lbs', and so on.<sup>7</sup> Still, I can think of many expressions which do not

<sup>6</sup>Throughout this discussion, 'right answer' must be understood somewhat narrowly. I am using 'right answer to Q' (where Q is a question) to cover a possible reply to Q if and only if a statement of the form 'A told B W' (where W must be thought of as replaced by the oratio obliqua of Q, and A and B by expressions through which two persons are mentioned) would be true of any episode in which the reply had been given by the person mentioned at A in response to the asking of Q by the person mentioned at B. This excludes such possibly "correct" answers as 'I don't know.' Certain questions do not admit of a right answer, such as, 'Who is the present king of France?'

<sup>7</sup>Significant distinctions exist among the expressions that I must reject. The first three would, under most circumstances, count as replies when uttered in response to an utterance of (4), but very special circumstances are called for in

belong in that lot, such as, '15,000 feet', '8,000 feet', '3 kilometers', and others.

I also take (5) to be a sound question: 'tea kettles do emit a humming noise just before the water in them begins to boil. I have views about its answer too: (a) it must be statable in the form of a sentence or conjunction of sentences preceded by the word 'because'; (b) this sentence or conjunction of sentences must include—perhaps among other things—a description of something that actually happens when water is about to begin boiling; and (c) this description must entail that on each such occasion something causes air to vibrate with an amplitude, frequency, and overtones corresponding to the loudness, pitch, and quality of the noise in question. As before, the conditions do not add up to an answer, even when supplemented with my other beliefs. Acceptance of them again requires that I reject a number of expressions as incapable of being formulations of the right answer. These too make up a varied lot: 'Because Cleopatra had a short nose', 'I don't know', 'Look it up in the encyclopedia', 'Ruth Otto', '19,321 feet'. But this time I am unable to conceive of any others—I can think of nothing, I can imagine nothing, I can conjure up nothing, I can invent nothing, I can remember nothing, that can survive confrontation with what I take to be conditions on the right answer. And this is clearly a consequence not only of the character of those conditions but also of the limits of my intellectual repertoire.

To recapitulate: (i) I take both (4) and (5) to be sound questions, to admit of a right answer.

(ii) I know, or believe I know, enough about each answer to be able to eliminate a number of possible utterances, that is, of expressions, as not being formulations of the answer.

(iii) In the case of (4), I can think of some possible utterances that I cannot eliminate in this way.

(iv) In the case of (5), I can think of no expression<sup>8</sup> that I cannot eliminate in this way.

the case of the last two. Furthermore, as a reply to (4) the first must convey a falsehood, the second may convey a truth, the third can convey neither a truth nor a falsehood. These distinctions are not relevant to our present purpose.

<sup>8</sup> "Can think of no expression" and "can think of no answer" (as short for "can imagine nothing, conjure up nothing, invent nothing, remember nothing, conceive nothing ...") do not cover the familiar states of momentary amnesia during which one has the answer "on the tip of one's tongue" but cannot utter it.

Let us say of anyone who stands with regard to any question *Q* in the relation in which I stand to (5) that he (she, they) is (are) in a *p*-predicament with regard to the question *Q*. In other words, *A* is in a *p*-predicament with regard to *Q* if and only if, on *A*'s views, *Q* admits of a right answer, but *A* can think of no answer to which, on *A*'s views, there are no decisive objections.

'*A* is in a *p*-predicament with regard to *Q*' is a neologism, but the meaning of each statement of that form is intimately related to the meaning of some statement of the form '*A* does not understand *q*', with *q* replaced by the oratio obliqua form of the question mentioned at *Q*. Any statement of the form '*I* am in a *p*-predicament with regard to *Q*' implies and is implied by the corresponding '*I* do not understand *q*'. In other words, the two kinds of statements are interchangeable when the person speaking is the person spoken about. The relation is a little more complicated when the speaker is not the person spoken about. In those cases '*A* is in a *p*-predicament with regard to *Q*' should convey what would be conveyed by an '*I* do not understand *q*' uttered by the person mentioned at *A*—with one difference to be specified in a moment—but need not convey what would be conveyed by the corresponding '*A* does not understand *q*'. This is a consequence of the following:

(a) Any '*A* does not understand *q*' implies that the question alluded to at *q* is sound. Thus its use by anyone conveys (pragmatically implies) that the user believes the question to be sound. Its cognate '*A* is in a *p*-predicament with regard to *Q*' does not imply that the question is sound, but only that it is sound on the views of the person mentioned at *A*; its use by anyone should therefore not convey that the user takes the question to be sound (except for a first-person *A*). In other words, the meaning of '*A* does not understand *q*' requires that the opinion of the user concerning the soundness of the question be dominant, but the meaning of '*A* is in a *p*-predicament with regard to *Q*' requires that the views of the person mentioned, about matters relevant to the soundness of the question, be dominant.

(b) An '*A* does not understand *q*' statement may be used to report or describe situations in which *either* (1) none of the answers that the person mentioned at *A* can conceive is an answer that that person can accept—and this includes situations in which the person spoken about can conceive of the right answer, but cannot accept it in the



light of his or her other beliefs—or (2) none of the answers that the person mentioned at A can conceive is the right answer—and this includes situations in which that person can conceive of one or more answers that he or she can consistently accept. These statements are thus marked by an ambiguity, and demand contextual clues that indicate whose conditions on the answer are at play, the speaker's or those of the person spoken about. 'A is in a p-predicament with regard to Q' suffers from no such ambiguity; it refers only to the sort of situations described under (1) above, that is, the meaning is such as to require that the relevant beliefs of the person mentioned at A be dominant. Here too the difference vanishes when the person spoken about is the person speaking.

An illustration will fix these two points. Neither Sam nor Rebecca know or think they know the answer to 'How did God make Eve, a whole human being, from just one bone?' Rebecca, a pure but simple soul, knows of no reason for dismissing the view that He first broke Adam's rib in as many pieces as He needed parts for her body, that He then stretched and shaped each of these into one of these parts, that He next assembled them into the well-known configuration, and that He finally blew life into the whole. Sam cannot accept this answer. He knows about the principle of conservation of matter, about the differences in chemical composition of muscles, hair, bones, skin, nerves, stomach, appendix. But he cannot think of anything that he would find acceptable. Max the atheist would reject the whole question. Sam would be right in saying that he himself is in a p-predicament with regard to the above question, and it would make sense for Max (or anyone else) to say that Sam is in a p-predicament with regard to the question, even though Max could not rightly say that Sam does not understand how God created Eve. But Sam, or Max, or Rebecca (or anyone else for that matter) would be mistaken in saying that Rebecca is in a p-predicament with regard to the question, since she can think of at least one possibility that she need not reject. Nevertheless, Sam would be correct in saying that she does not understand how God created Eve. No one should say that Max is in a p-predicament with regard to the question.

There is one respect in which an 'A is in a p-predicament with regard to Q' statement not made by the person mentioned at A conveys less than its cognate 'I do not understand q' when uttered

by that person. The latter statement conveys that that person knows or believes that he or she does not understand q, but the former does not entail that the subject knows or believes that he or she is in a p-predicament with regard to Q, and could not convey it.

Being in a p-predicament with regard to a question may sound like greater ignorance than merely not knowing the answer. It isn't. Being in a p-predicament usually calls for more learning than does mere ignorance. Sam would not be in a p-predicament with regard to 'How did God make Eve?' had he not studied chemistry. It was once suggested to me that the noise emitted by tea kettles just before the water begins to boil might come from surface vibrations set off by air bubbles released before vapor bubbles begin to form. For a while thereafter I was not in a p-predicament with regard to 'Why do tea kettles emit a humming noise just before the water begins to boil?' Since then I have had occasion to observe the phenomenon a little more closely, and I have satisfied myself that these air bubbles are too small and too few to account for the noise in question. I would not be in a p-predicament today had I not learned these facts. Or again, I am not in a p-predicament with regard to 'What is the relation between the Debye temperature  $\theta$  and the temperature at which the atomic heat of a solid element is equal to  $3R/2$ ?' I hardly know enough to understand the question; the only condition I associate with it is that the answer must be statable as a mathematical function of  $\theta$ , but I can construct indefinitely many of those!<sup>9</sup>

*Second hypothesis:* A statement of the form 'A explained to B W' is true if and only if an episode has occurred

<sup>9</sup>All my examples have been homespun cases. To lend dignity and significance to this discussion it would no doubt be better to use illustrations taken from the history of science. I have not done this because the analysis required to show that some scientist was in a p-predicament with regard to some question at some time would be very long and very complex: it would demand an examination of theoretical background, available evidence, implicit assumptions, and the like. But it is obvious that there have been times in the history of science when scientists were in a p-predicament with regard to some question, for example, when a maximum height to which water can be pumped was known but the concept of atmospheric pressure had not yet been discovered; when the mathematical representation of energy distribution for black body radiators was known but the quantum hypothesis had not been thought of or admitted; when the Michelson and Morley results were known but the Special Theory of Relativity had not been advanced, and so on.

- (a) at the time of which the tutor knows the right answer to the question;<sup>10</sup>
- (b) at the beginning of which the tutee was in a p-predicament with regard to the question;
- (c) at the beginning of which the tutee thought of himself as in a p-predicament with regard to the question;
- (d) during which the tutor knows, or believes, or at least assumes that, at the beginning of the episode, the tutee was in a p-predicament with regard to the question;<sup>11</sup>
- (e) in the course of which the tutor presents the facts that in his opinion the tutee must learn to know the right answer to the question;
- (f) at the end of which all the facts mentioned in (e) have been presented to the tutee by the tutor;
- (g) at the end of which the tutee knows the right answer, and knows it as a result of, and in virtue of what he has been told by the tutor in the course of the episode.

Some of these clauses obviously do not stand for necessary conditions. (b) and (g) would require initial ignorance and ultimate knowledge of the right answer by the tutee. This is uncalled for, as we saw in the comments to the first hypothesis. But (c) and (d) are not necessary either. (c) too implies initial ignorance of the right answer. (d) (and incidentally (c)) would imply that the following correct account is incorrect. Until the age of five I was convinced that babies grew in cabbages. My older brother knew this about me, and on my fifth birthday, as a present, he told me the facts of life; on that day he *explained* to me how babies come into the world. At the age of four, I was not, nor did I think myself to be,

<sup>10</sup>This implies of course that the question must have a right answer.

<sup>11</sup>'Assumes' throughout this discussion must carry a heavy load. To assume may simply be to take for granted. It may be something else. When I give an examination to my students, I always tell them to assume that they are writing for someone who knows nothing about the subject; but I do not want them to take for granted that I know nothing about the subject. I always assume that they know very little philosophy when I prepare a lecture for them, though I often have no strong convictions on this. 'To assume' is thus sometimes 'to go on the assumption that'. It can also be 'to take for granted' and 'to go on the assumption that'. In stating these hypotheses, I have assumed that 'to assume' may do the job of a disjunction of all these possibilities.

in a p-predicament with regard to 'How do babies come into the world?' since I had an answer, though admittedly a false one. My brother knew this about me, and the statement that he explained to me how babies come into the world does not imply the contrary, that is, it does not imply that he thought or assumed that I would not have been able to come up with an answer compatible with my total set of beliefs. This is in general true of statements of the form 'A explained to B W'. They do not imply that the tutor (or tutee) believed or assumed that the tutee would have been or should have been puzzled (in a p-predicament way) by the question before the explaining episode.

If we eliminate all these conditions, however, we are left with (a), (e)—'learn' being perhaps replaced by 'learn or know'—and (f), a set enough like the revised first hypothesis to be recognized immediately as not sufficient.

On the other hand, this hypothesis in its original form, unlike the first one, does seem to provide a set of sufficient conditions. I don't know how one can prove this. My own conviction on the matter rests on the fact that I am unable to construct any illustration that meets these seven clauses yet is not an explaining episode. It helps, in this connection, to look at the counterexamples showing the insufficiency of the previous hypothesis.

The first counterexample, it will be recalled, consisted of an episode in which I was told the height of the Empire State Building, something I had not known before, but knew thereafter. But 'What is the height of the Empire State Building?' is not a question with regard to which I was or could have been believed to be in a p-predicament (the same, by the way, is true of 'What is the height of Mt. Kilimanjaro?').

The second example was an episode in which a prisoner told how he had disposed of the tunneled dirt, and, in effect, settled in the mind of his audience which of a number of obvious methods he had actually used. Here again the tutee, that is, his audience, was neither known nor thought to be in a p-predicament with regard to the question. This illustration was contrasted at the time with another one, in many respects similar, but which was clearly recognizable as an explaining episode, that is, as properly describable by an 'A explained to B W' statement. The latter episode, notice, does satisfy the present hypothesis.



The present hypothesis has a minor virtue that may strengthen its *prima facie* plausibility as a set of sufficient conditions. The etymology of a word, we have frequently been told, often provides the key to the analysis of its meaning. The origins of 'to explain' and of its French cousin *expliquer* go back to expressions used to speak of making smooth by removing folds and wrinkles. But what were these ancestral folds and wrinkles? The literature contains a number of suggestions. In *La Théorie Physique*, Pierre Duhem offers the following one: "To explain [*expliquer*], *explicare*, is to strip reality of the appearances in which it is wrapped as in veils, in order to see this reality naked and face to face." There is a more Lockean<sup>12</sup> and more plausible possibility compatible with a more modest reality. The second hypothesis would support it. People who contemplate a question with regard to which they are in a p-predicament are prone to frown their foreheads, to screw up their faces, to knit their brows, and they usually shed most of these folds and wrinkles and present a smoother countenance upon being told the answer. We know that one who remedies a p-predicament often explains in this sense which, if original, is now obsolete. Is this account correct? Perhaps not. (And yet ... the first quotation under 'to explain' in the O.E.D. is: "He must caulk and explain his forehead," 1569; and the second reads "Their faces are explained and flattered by art," 1650.) For our purpose, that does not really matter. It is enough if the account reminds us that sentences of the form 'A explained W to B' are aptly chosen to report episodes in which a tutor turns someone who could truly have said 'I don't understand W' into someone in a position to assert 'I know W', and if it provides the occasion for seeing that the hypothesis does justice to this fact.

All this suggests that the clauses making up the second hypothesis should perhaps be divided into three classes: first, some that are necessary conditions, (a), (e), and (f); second, some that are totally dispensable, (b), (c), and (g); and third, one that is not necessary by

<sup>12</sup> "It may also lead us a little towards the original of all our notions and knowledge, if we remark how great a dependence our words have on common sensible ideas: and how those which are made use of to stand for actions and notions quite removed from sense, have their rise from thence, and from obvious sensible ideas are transferred to more abstruse significations, and made to stand for ideas that come not under the cognizance of our senses; ..." John Locke, *Essay Concerning Human Understanding*, Book III, chap. 1, sec. 5.

itself but is a member of a set of conditions whose *disjunction* is a necessary condition, (d).

At the beginning of this discussion I pointed out that a good analysis of explaining episodes should show why some indirect questions are out of place as direct objects of 'to explain'. An account could be built on the second hypothesis. That account would not rest on any of the "totally dispensable" clauses listed above, and may therefore be part of the truth. To the extent to which it strikes one as such, it will confirm my suggestion that the present hypothesis is a disjunct of the truth that we are after. I will therefore give a rough sketch of that account.

*Indirect questions beginning with 'whether'.* These represent questions whose right answers must be either 'yes' or 'no'. But no tutee can be thought to be in a p-predicament with regard to such a question. If the views held by him require that he reject both 'yes' and 'no' as each violating some belief or assumption of his, then he cannot take the question to be sound, to admit of a right answer. But then he cannot be in a p-predicament with regard to the question either, since one can only be in a p-predicament with regard to questions that one can consistently think of as having a right answer. On the other hand, if he is not compelled to reject both 'yes' and 'no', he cannot be in a p-predicament either. Could a tutee be in a situation in which he can think of 'yes' (and have to reject it) but not of 'no', or vice versa? Perhaps, but then the question is one that he does not understand, and indefinitely many answers will be open to him. In the case of these indirect questions, then, nothing corresponds to the sort of knowledge or belief or assumption called for in (d) and the conditions of the hypothesis cannot be satisfied.

*Indirect questions beginning with, for example, 'What is the height of ...?', 'What is the weight of ...?', 'What is the distance between ...?', that is, questions calling for an answer expressible in the form of a number and a unit of measurement.* A tutee who knows or thinks he knows the form of the answer and who knows or thinks he knows appropriate units of measurement must either reject every number, in which case he cannot take the question to be sound, or there is some number that he does not have to reject. But how could he be in a p-predicament in the latter case? To be able to conjure up further answers he need only be able to generate numbers. One can imagine situations in which a tutor believes that he

is dealing with a partly educated tutee unable to count, yet able to reject the possibilities that would occur to him. But in those cases, if the conditions of the hypothesis are satisfied, an 'A explained to B what ...' would be true. And in the others (d) could not be satisfied. In the remaining cases, where the tutee does not think that he knows the form of the answer, he is in a situation analogous to the one that I described as mine with regard to the question about the Debye temperature, and thus not in a p-predicament because in no position to reject possibilities apt to occur to him. If the tutor believes that this is the situation, (d) is not satisfied. One can also imagine cases where the tutee is thought of as in a position to reject every kind of unit of which he is aware but as not aware of the right kind of unit. In those cases, however, 'A explained to B W' statements will be true, if the conditions of the hypothesis are all satisfied.

*Indirect questions beginning with 'which'.* These represent interrogatives that open with the word 'which'. An interrogative sentence beginning with the word 'which' is not the formulation of a question unless and until it becomes associated with a specific set of alternatives that include the right answer, if any. Anyone thought to be in a position to have views on the soundness of the question must also be thought to know the members of that set. But then (d) cannot be satisfied. The tutor must either believe that the tutee should reject every member of the relevant set and ipso facto should reject the question itself as without correct answer, or the tutor must believe that one alternative at least is an open possibility on the views of the tutee. Either belief rules out thinking that the tutee is in a p-predicament with regard to that question.

### B-predicament

The third hypothesis too requires the introduction of a new phrase, 'being in a b-predicament with regard to a question'. A statement of the form 'A is in a b-predicament with regard to Q' (in which A must be replaced by an expression referring to some person or persons, and Q by one referring to some question) is true if and only if the question mentioned in it admits of a right answer, but that answer is beyond what the person mentioned can conceive, can think of, can imagine, that is, is something that that person cannot remember, cannot excogitate, cannot compose.

My position with regard to 'Why do tea kettles emit a humming

noise just before the water begins to boil?' can be used to bring out the difference between a p-predicament and a b-predicament. I am in a p-predicament with regard to that question. I am also in a b-predicament with regard to it, since I am unable to think of the right answer. However there was a time when I was in a b-predicament with regard to the question *without* being at the same time in a p-predicament with regard to it, when I was able to entertain 'because the surface of the water is made to vibrate by escaping air bubbles that form before vapor bubbles do' as possibly the correct answer. Note that it is also possible to be in a p-predicament with regard to a question *without* at the same time being in a b-predicament with regard to it. This is the case when someone is in a p-predicament with regard to a question that in fact admits of no right answer. It is also the case when someone rejects the correct answer to a question with regard to which he is in a p-predicament. Such would be my situation if the air-bubble theory were correct. Such was Kepler's situation with regard to 'What is the shape of planetary orbits?' while he refused to admit the possibility of elliptical paths.<sup>13</sup>

*Third hypothesis:* A statement of the form 'A explained to B W' is true if and only if an episode has occurred

- (a) at the time of which the tutor knows the right answer to the question;
- (b) during which the tutor knows, or believes, or at least assumes that at the beginning of the episode, the tutee was in a b-predicament;
- (c) in the course of which the tutor presents the facts that, in his opinion, the tutee must learn to know the right answer to the question;
- (d) at the end of which all the facts mentioned in (c) have been presented to the tutee by the tutor.

<sup>13</sup>'A is in a b-predicament with regard to Q' is in some respects closer to 'A does not understand W' than is 'A is in a p-predicament with regard to Q'. It requires a sound question. Its truth is independent of the opinions held by the person mentioned concerning the soundness of the question or concerning the conditions that the right answer must satisfy. Notice, however, that 'A is not in a b-predicament with regard to Q' unlike 'A understands W' does not require that the question be sound.

This hypothesis includes or implies none of the clauses shown so far to stand for conditions that are not necessary conditions. It does not require initial ignorance or ultimate knowledge of the answer by the tutee, and it does not require that the tutor think the tutee to have been in a p-predicament at the beginning of the episode.<sup>14</sup>

Like the second hypothesis, the third seems to be a set of sufficient conditions. Again, I can't prove this, but again I am unable to construct a case that meets the requirements of the hypothesis but is not an explaining episode. Those that brought out insufficiency before won't do so here. I was not nor could I have been believed to be in a b-predicament with regard to 'What is the height of the Empire State Building?' Our prisoner while the protagonist in a nonexplaining episode was addressing an audience that was not, and was not thought to be, in a b-predicament with regard to the question. But his audience was believed to be in a b-predicament (as well as a p-predicament) in the explaining episode. Furthermore, the illustration showing that (d) of the second hypothesis is not a necessary condition, satisfies the conditions of the present hypothesis: my brother knew that I was in a b-predicament with regard to 'How do babies come into the world?'

Speculative etymology may stir the relevant intuitions here too. The prototypic folds should now be brought to mind by the story of the excitable Frenchman who, when shown an oldfashioned telephone requiring the earpiece to be held in one hand and the mouthpiece in the other, protested: "Mais alors, comment s'explique-t-on?" Words often seem too weak to cure a b-predicament, gestures seem called for, unfoldings of arms, openings of hands. These may have been the original topic of 'A E to B W' descriptions! The O.E.D. fails us at this point, but not completely; under 'To explicate' we find "When he intendeth his business to purpose, then he standeth upon his feet, explicateth and displayeth his limbs" (1620).

This hypothesis, too, will support an account of why some indirect questions are out of place as direct object of 'to explain'. Its

<sup>14</sup>Someone in a b-predicament with regard to a question cannot know the answer to that question, but (b) is phrased to allow for explaining episodes at the beginning of which tutees do know the answer. (c) and (d) furthermore are phrased to allow for explaining episodes in which tutors exert themselves in vain, are not understood, are not listened to attentively, etc., that is, explaining episodes at the end of which the tutee still does not know the answer.

outlines are readily drawn; the details would get us into the arid territory of erotetic logic.

*Indirect questions beginning with 'whether'.* To believe that someone is in a b-predicament with regard to a question whose right answer must be either 'yes' or 'no' is to believe that someone is incapable of thinking of either 'yes' or 'no'. This is not merely to believe or to assume that the person is unable to accept, or to conceive of himself as accepting either 'yes' or 'no' as the correct answer; it is in effect to impute to that person the inability to come up with the thought of 'yes' or 'no' as even the wrong answer. On the present hypothesis, an 'A E to B whether ...' statement cannot be true unless some tutor (i) believed or assumed some tutee to be that colossally dull and yet (ii) undertook to convey information to that tutee. But (i) and (ii) are pragmatically incompatible: the occurrence of (ii) would establish the falsehood of (i). Thus the conditions of the hypothesis cannot be met in these cases.

*Indirect questions like 'What is the height of ... ?' 'What is the weight of ... ?', that is, indirect questions calling for a number followed by a unit of measurement.* The cases involving units of measurement familiar to practically everyone can be dealt with in the same way as indirect questions beginning with 'whether'. The fact that a tutor undertakes to inform a tutee is prima facie evidence that he does not believe or assume that one specific combination of a number and an obvious unit of measurement surpasses what the tutee can conceive. Normally the conditions of the hypothesis are not fulfilled.

The cases involving less familiar dimensions and less familiar units are a little more complicated. Indirect questions standing for questions of that sort *can* serve as direct object of 'to explain'. Their aptness in that role depends, however, on how much the tutor tells the tutee. Thus

- (6) Picard explained to Mme. Rouge what the difference of potential between the poles of his car battery is.

is not warranted if Picard merely said to Mme. Rouge, "The difference of potential between the two poles of my battery is six volts." But then that would be prima facie evidence that (b), (c), and (d) are not satisfied. But (6) is *not* unwarranted if Picard presented Mme. Rouge with a lengthier exposition, one designed to convey

not only the sound of the answer, but its meaning as well, that is, designed to put her in possession of notions not previously in her intellectual repertoire. This too is compatible with the hypothesis.

*Indirect questions beginning with 'which'.* To hold that a tutee is in a b-predicament with regard to a question beginning with 'which' a tutor would have to assume that the tutee associates with it the same set of alternatives that he does: these alternatives are part of the very identity of the question. The tutor would also have to assume that one of those alternatives is the right answer. It is therefore impossible for a tutor to believe or to assume that some tutee is in a b-predicament with regard to such a question.

Even so, the third hypothesis cannot be a set of necessary conditions. Consider the following example:

I am the Dr. Watson to some Sherlock Holmes. Yesterday morning we were called to investigate a crime. The body was found in a room with a fireplace, two windows, one door. According to irrefutable evidence, the doors and the windows were locked from the inside, by the victim, before he was killed, and were not touched again until investigators entered the room. The chimney is too narrow to allow the passage of a grown man. Walls, floor, and ceiling have been checked; there are no secret doors, no hidden outlets. The victim did not commit suicide: he was killed by a bullet in the head, was not wearing gloves, and the fingerprints on the gun do not match his.

How did the assassin leave the room?

All day yesterday I was in a p-predicament with regard to that question. The only answers that occurred to me were ones that I was forced to reject: through the door, the north window, the south window, the chimney, the floor, the ceiling, the walls.

At dinner time, my Sherlock Holmes, who knew my condition, explained to me how the assassin had left the room: through the chimney. He was a midget, small enough and slender enough to get through it.

This counterexample does not come as a surprise. The third hypothesis cannot be a set of necessary conditions if the second hypothesis is a set of sufficient ones, since these two hypotheses can be independently satisfied. The midget case is merely an explaining episode that accords with the second but not with the third. But it confirms the hunch that the final doctrine must contain a disjunctive clause.

*Fourth and Last Hypothesis:* The essential characteristics of explaining episodes are the following:

- (a) the question is sound, that is, admits of a right answer;
- (b) the tutor is rational and knows the right answer to the question at the time of the episode;
- (c) during the episode the tutor knows, or believes, or at least assumes that at the beginning of the episode the tutee was in a p-predicament with regard to the question, or that at the beginning of the episode the tutee was in a b-predicament with regard to the question, or that at the beginning of the episode the tutee was in either a p-predicament or b-predicament with regard to the question;
- (d) in the course of the episode the tutor presents the facts that in his opinion the tutee must learn to know the right answer to the question;
- (e) in the course of the episode the tutor also provides the tutee with such instruction as he (the tutor) thinks necessary to remove the basis of whichever of the states mentioned in (c) he deems the tutee to be in;
- (f) at the end of the episode all the facts mentioned in (d) and (e) have been presented to the tutee by the tutor.

I have expressed this hypothesis in the material mode, and have included three clauses, that is, (a) and (c) and (f), that are superfluous since (a) is implied by (b), and (c) and (f) by (d) and (e). I have chosen the material mode because the hypothesis concerns more than our understanding of an English verb.<sup>15</sup> The superfluous clauses are apt to be overlooked and deserve an explicit statement.

A word about (e). It excludes from the class of explaining

<sup>15</sup>There is another reason for using the material mode. Put in the formal mode, this hypothesis provides only truth-conditions for statements in the form 'A explained to B W', that is, with 'to explain' in the simple past tense. But we are interested in all 'A E to B W' statements, regardless of the tense of the verb. Extending the above conditions to include cases with other tenses would require the discussion and resolution of difficult problems about the relationship among truth-conditions of statements in every respect alike but for the tense of the verb. The use of the material mode should enable us to visualize the relevant points in a way that is adequate for present purposes; for example, the continuous present is applicable while an explaining episode is going on, the past tenses if it took place in the past, the future tenses if it is to take place, and so on.

episodes those episodes in which the tutor provides no more information than might have been put in the form 'The answer to Q is ...', for instance, 'The answer to "How did the assassin leave the room?" is "through the chimney".' In one sense, such statements, if true, contain all that one has to learn to know the answer to the question. But they cannot make up the total substance of explaining episodes.

It is easier to indicate what (e) excludes than to indicate what it is designed to include. A fully adequate treatment of what is involved would take us far afield. I shall evoke the relevant facts through illustrations.

Picard, in a previous example, it will be recalled, knew or believed, or at least assumed, that Mme. Rouge was in a b-predicament with regard to the question 'What is the potential difference between the two poles of Picard's car battery?' How could he know, or believe, or assume this? He knew that the answer has to be expressed in volts. He also knew that 'volt' is what some have called a 'theory laden term', that is, he knew that before one can understand the term 'volt' one must become aware of certain phenomena and of principles, concepts, ideas, and facts used to represent, to reason about, and to record these phenomena. He furthermore knew or believed, or at least assumed that Mme. Rouge did not possess the background just alluded to, that she had never acquired the modicum of factual and theoretical knowledge required to understand 'volt'. He assumed that she could not have conjured up the answer because he assumed that she literally did not have the concepts that she would have had to have. The basis of her alleged b-predicament with regard to the question presumably went deeper than mere ignorance of the words needed to formulate the answer; to remove it Picard had to enlarge her conceptual repertoire, that is, he had to teach her some elements of electricity.

Removal of the basis of a b-predicament or p-predicament with regard to a question will in general require instructions about matters not covered by a mere statement of the answer, but the above example ought not to mislead us into thinking that it always requires *theoretical* instruction. Remember how my brother, on my fifth birthday, removed the basis of my b-predicament. Even less theory will often do. Picard does not know who the author of *De Fabrica* was. He is in no position to think of the answer because he

has never heard of Vesalius. To remove the basis of his b-predicament with regard to 'Who was the author of *De Fabrica*?' one must do more than give him the name of the author; to explain to him who the author of *De Fabrica* was one must tell him who Vesalius was. But this demands no abstract principles or theoretical notions. Or consider the case of someone assumed to be in a p-predicament but *not* in a b-predicament with regard to some question. I was in a p-predicament with regard to 'How did the assassin leave the room?' because I took it for granted that the culprit was a full-sized, normal human being. The correct answer was ruled out for me by a false assumption. In all such cases removal of the basis of the p-predicament requires instructions that lead one to *revise* one's views. (Einstein had to do something of the sort when he explained why no interference pattern had been observed by Michelson and Morley. In his episode theory had to be expounded, but not so in that of Holmes.) Holmes would not have explained how the assassin had left the room had he merely said 'through the chimney', but the additional information was short and pedestrian.

What is to be said in behalf of this fourth hypothesis?

If the second hypothesis (with its "totally dispensable" clauses removed) and the third hypothesis are each a set of sufficient conditions, then this one must also be a set of sufficient conditions. This hypothesis does not contain any of the conditions, either explicitly or implicitly, that we have had occasion to reject as not being necessary ones. It includes all the conditions that we have had occasion to recognize as necessary ones. It lends itself to an account of why some indirect questions are out of place as direct objects of 'to explain'. (The account is readily constructed by combining those suggested by the second and third hypotheses.) It even lends itself to the speculative etymology to which the second and third ones lend themselves. I, for one, cannot think of an episode that fits all the clauses in this hypothesis and that is not correctly described by an 'A explained to B W' sentence, nor can I imagine an episode that is so describable but that does not fit all the clauses of this hypothesis.

There is one possible objection that ought to be considered. We speak on occasion of someone's having tried to explain something to someone without having succeeded, and we often imply thereby that at the end of the episode the tutee did not know the answer. This may suggest that ultimate knowledge is necessary after all, since to

have tried to explain something to someone without succeeding is no more to have explained something than to have tried to scale Mt. Everest without succeeding is to have scaled it.

But the fourth hypothesis accommodates these utterances in a natural way. Statements of the form 'A tried to explain to B W but failed' cover unfinished explaining episodes, that is, episodes in which (f) is not satisfied, and which therefore do not meet the truth-conditions proposed here. They cover episodes in which a tutor comes to realize that he is not being listened to and gives up; they cover episodes in which he is prevented from finishing; they also cover episodes in which he withholds nothing yet realizes that there are some things, things that have not occurred to him, that the tutee must be told to become properly informed under (e). Notice that they do *not* cover episodes in which all the conditions of our hypothesis have been satisfied, but at the end of which the tutee is still ignorant because of *his* failure to attend properly.

All this does not amount to a proof of the fourth hypothesis, but it goes a long way toward establishing a presumption in its favor.

### 3

I now turn to statements of the form 'A E X to B' and 'A E to B X' where A and B indicate positions occupied by expressions through which some person or persons are mentioned, where E indicates a position occupied by a tensed version of 'to explain', but where X is *not* replaced by an indirect question.

Statements of that form also describe, or report, or record the occurrence of explaining episodes. 'To explain' in each functions as an accomplishment term, and their truth entails that the truth-conditions of some 'A E to B W' statement (with identical replacements for A, B and E) are satisfied. Their own truth-conditions should therefore be readily obtainable from those of 'A E to B W' through some appropriate modifications.

Where X is replaced by 'something' (for example, 'Sam explained something to Rebecca') the question is simply left unspecified, and we can adapt the fourth hypothesis by substituting for (a),

- (a<sub>1</sub>) There is a sound question, that is, one that admits of a right answer.

The remaining cases fall under two headings, those in which X

is replaced by a noun phrase (for example, 'Hector's behavior', 'the laws of electrostatics') and those in which it is replaced by a statement in oratio obliqua form (for example, 'that bones contain no dextrose').

First, those in which a noun phrase occurs. In 'A E to B W' statements the question is specified by an indirect question. However, questions can be specified in other ways and some noun phrases in some contexts can fulfill that function and can serve as surrogates of indirect questions. Thus 'Hector's behavior' will stand for 'why Hector behaves (or behaved) in the way he does (or did)' and 'the origins of the word "explain"' will stand for what the origins of the word 'explain' are. The precise description of the relevant transformations is something that we must leave to grammarians. The fourth hypothesis will cover statements with such noun phrases if we amend the definition of 'the question' on page 22 to read "the question whose oratio obliqua form or a surrogate for whose oratio obliqua form would appear at the position indicated by W or X."

Some occurrences of noun phrases do not convey enough information to be considered as surrogates of indirect questions. Mere knowledge of English is not sufficient to produce the question in an episode described by 'Mme. Rouge explained the third chapter of *Moby Dick* to Picard'; one must be acquainted with the third chapter of *Moby Dick* and know something of the questions to which it gives rise. The vagueness characteristic of such statements must be preserved in an adequate formulation of their truth-conditions. But this is easy to achieve. Simply replace (a) of the fourth hypothesis by

- (a<sub>2</sub>) There is a question about the topic mentioned at the X position that is sound, that is, that admits of a right answer.

Notice that though such noun phrases are vague, they are nevertheless indispensable. For instance, the third chapter of *Moby Dick* gives rise to a cluster of questions. It can happen that a tutee is not in a p-predicament with regard to any one of these questions taken in isolation, and that he is able to admit each as sound, but that he is nevertheless not in a position to think of a *set* of answers (to that set of questions) that he can consistently admit as possibly correct. This will happen when every nonobjectionable answer that he can think of to one isolated question would require him to reject



every nonobjectionable answer that he can think of to another isolated question. Even more complicated situations arise. A question with regard to which a given tutee is in a p-predicament may itself rest on a presupposition that the tutee trusts simply because it is the only nonobjectionable answer that he can think of to some other question. Giving him the correct answer to the latter may teach him that the former has *no* right answer, and may thus "cure" him of his p-predicament. In all such situations "the question" is complex and is made up of disjunctions and conjunctions of more elementary ones. In each case it can no doubt be put in oratio obliqua form but the result must be unwieldy and awkward within an 'A E to B W' statement.<sup>16</sup>

This point is worth remembering. Often the question in an explaining episode cannot be told from an 'A E X to B' description of that episode. This may lead us to overlook the essential connection that links explaining to questions.

Now, briefly, the cases in which the direct object of 'to explain' is a statement in oratio obliqua. The nature of these statements is obvious enough: they reproduce, if not the words, at least the gist of what the tutee was told. To pass from a description containing such a statement to a description containing a question in oratio obliqua one merely has to phrase the question calling for the statement. This is sometimes easy. But not always. A given statement may convey the answer to more than one question. Moreover, the alluded-to words may have conveyed not the content of the answer but part of the background. Some explaining episodes, as we know, require that the tutee be corrected about some assumption, and others that he be informed of matters that will put the answer within his intellectual grasp. To know the role played by something that the tutor told the tutee one must know more than what he said. This will be reflected in a statement of truth-conditions obtained by substituting ( $a_1$ ) again for (a). "There is a sound question, that is, one that admits of a right answer"; and by supplementing (f) to read "At the end of the episode all the facts mentioned in (d) and (e) have

<sup>16</sup>To simplify the discussion I have written as if the actual cognitive state of the tutee were essential rather than what the tutor believes or assumes about that state, and I have ignored b-predicaments altogether. The necessary adjustments are obvious enough, and do not merit the further circumlocution that their exposition would require.

been presented to the tutee by the tutor, and these include the facts described by the statement at the direct object position."

## 4

I pass now to statements in which 'to explain' does not function as an accomplishment term. First, those of the form 'A explained W' in which A still indicates a position occupied by an expression through which a person is mentioned. When one says that Newton explained why there are tides, one need not mean that some explaining episode took place in which Newton was the tutor. One may mean that Newton solved the problem, found the answer to the question. Similarly 'Sherlock Holmes explained how the assassin left the room' may convey that Holmes solved the riddle. Here 'to explain' functions as an achievement term, not as an accomplishment term.

The fourth hypothesis does not require but is compatible with the occurrence of explaining episodes at the beginning of which the tutee is really in either a p-predicament or b-predicament with regard to the question, and by the end of which he actually has learned the answer and has had the grounds of his predicament corrected. I shall call explaining episodes of that sort *proper explaining episodes*.

A proper explaining episode involves a change in the tutee: at the beginning he does not know the answer to the question; in the course of the episode he learns certain facts, or (vel) he gains certain ideas and concepts, or (vel) he becomes aware of certain principles; at the end of the episode, he knows the answer, and he is competent to act as tutor in explaining episodes whose tutees are in a p-predicament or b-predicament with regard to the same question—they do not know some or all of the facts, ideas, concepts, or principles that our initial tutee has acquired in our initial episode. In other words, he is changed into someone who can explain the matter.

The kind of transformation just described can be brought about without benefit of tutor and without the occurrence of an explaining episode: research, reexamination of beliefs and assumptions, conceptual inventions, explorations, prodding for inspiration, exploitation of good luck, or grace may enable one to transform himself by himself into somebody able to explain something. To have succeeded in this sort of endeavor is to have explained something in the sense of 'to explain' to which we are now attending.

The connection, then, between the truth-conditions of 'A E W' and 'A E to B W'—put in the material mode—is that to have explained something in the sense now under consideration is to have become able to explain something in the sense discussed in the fourth hypothesis, as a result of one's own endeavors and ingenuity.<sup>17</sup>

## 5

Now finally the statements of the form 'T explains W' in which T is *not* a position occupied by an expression through which persons are mentioned. For instance, 'The Special Theory of Relativity explains why no interference pattern was observed by Michelson and Morley', 'Lumumba's prison record explains why he hated everything Belgian', 'The fact that the assassin was a midget explains how he was able to leave the room without touching door or window'. In such statements 'to explain' functions as a state term.

One might expect a characterization of the truth-conditions of these statements to require a list of the kinds of references that may be made at T. But this would draw us into a veritable philosophic Sargasso Sea: theories, facts, events, propositions, items of information, discoveries, principles.... The list seems to have no end, and each of its entries may ensnare and choke us. Fortunately, no such list is needed.

The transformation of someone into one who can explain something does not always include an explaining episode. We noted in the previous section that one can achieve this by oneself. The change can often be initiated by informing the tutee of things that will put him in a position to work out the answer and to correct by himself the deficiencies responsible for his predicament. About any question W with regard to which someone might be in either a p-predicament or a b-predicament, one may ask: *What must someone who is in a p-predicament or a b-predicament with regard to that question learn,*

<sup>17</sup>As an achievement term 'to explain' is also often used to credit people with certain scientific discoveries. 'Newton explained why the tides vary with the phases of the moon' may serve to mark the fact that Newton was the one who solved the riddle, who found the answer to a question with regard to which everybody had been in either a p-predicament or a b-predicament. To have explained something in this sense is to be one of the first to have explained it in the sense just analyzed in section 4, that is, to be one of the first to have been in a position to explain (in the sense of section 2) it to a tutee.

*or of what must he be made aware to become able to explain W?* Any noun phrase that is the right answer to *that* question may appear as the subject of a statement of the form 'T explains W'. Furthermore, *such a statement is true if and only if the subject at T is related to the question represented at W in the way that a right answer to a 'What must someone in a p-predicament or a b-predicament with regard to W learn, and of what must he be made aware to become able to explain W?' question is related to the question at W.*<sup>18</sup>

This question-using way of putting things is clearly a question-begging one. What is the relation referred to in the above bi-conditional? Under what conditions does it hold? One wants here a set of standards, a set of criteria, and an explicit general procedure that determine whether something that purports to explain something does indeed do so. This I cannot give for the present. Getting what is wanted here must wait for developments in at least two areas: (a) erotetic logic, and (b) the theory of p- and b-predicaments. Notice, by the way, that the latter is not to be thought of as the concern of psychology. Situations in which people are, or are not, in a p- or a b-predicament with regard to some questions depend on logical factors first, and on psychological ones only incidentally. For instance, some theories generate questions that, according to other theories, rest on unsound presuppositions. That this is so, is in each case, a matter of logic. It follows, however, that one who subscribes to one of the former theories can be in a p-predicament with regard to questions with regard to which someone who subscribes to one of the latter theories will not be in a p-predicament. Or again, two systems of assumptions and of concepts may differ in that only one is capable of generating answers to questions to which both provide presuppositions. This again must be a matter of logic. But it follows that one who relies on a specific system of assumptions may necessarily be in a p-predicament or b-predicament with regard to a question to which someone else, aware of different systems, will be able to envisage answers, perhaps even correct answers. The object of a theory of p- and b-predicaments must be to provide a general and systematic treatment of all such possibilities and of their grounds.

<sup>18</sup>The use of the past tense of 'to explain' here is worth noticing. It implies that the conditions for the above relation do *not* hold though they were once thought to hold.

One last point before we leave 'to explain'. Everyone in a p-predicament or b-predicament with regard to a specific question does not always need the same instruction. A may be in a p-predicament with regard to the question as a consequence of one set of objections to the right answer; B may not share those particular objections but may have some of his own; C may not be in a p-predicament at all but only in a b-predicament; D may be in both. Though the question is the same, in each case a proper explaining episode may have to correct different shortcomings and inform on different points. Therefore the truth of statements of the form 'T E W' is relative to something left unmentioned by them, that is, the basis of the p-predicament or b-predicament with regard to the question. Such statements taken out of context may thus suffer from a peculiar ambiguity: they are true when viewed against such and such nescience, but not when viewed against some other.

## 6

What is an explanation? An explanation may be something about which it makes sense to ask: How long did it take? Was it interrupted at any point? Who gave it? When? Where? What were the exact words used? For whose benefit was it given?

It may also be something about which none of these questions make sense, but about which it makes sense to ask: Does anyone know it? Who thought of it first? Is it very complicated?

We are not dealing here with two kinds of explanation, but with two senses of the word 'explanation', one that refers to a certain type of didactic performance, and instances of it, the other that refers to something more abstract, to something that constitutes the cognitive substance of such performances. I will therefore speak of a *performance sense* of 'explanation' and of a *text sense* of 'explanation'.

Each instance of *the* explanation (in the performance sense) of something coincides with an explaining episode; the two form necessarily coextensive classes. Thus the explanation of something by someone is merely the performance by someone of the role of tutor in an explaining episode.

Instances of *an* explanation (in the performance sense) of something form a wider class. They cover not only genuine explaining episodes, but also episodes corresponding to what I have called a

looser use of 'to explain', that is, episodes in which the tutor *thought* that he was explaining something, but was mistaken and did not know the answer to the question, or did not give it.

The text sense of 'explanation' is more elusive. Ontological snares cover the genus within which its reference subsists, and should discourage a straightforward extensional explication. Something is *the* explanation (in the text sense) of something if and only if it is the right answer to a question formed from:

- (7) What must one tell someone in a p-predicament or a b-predicament or either a p-predicament or b-predicament with regard to Q in order to explain (in the accomplishment sense) W to him?

by putting a question at the position indicated by Q and the oratio obliqua form of that question at the position indicated by W, the resulting question being understood as calling for an actual utterance of what must be said to the tutee, not a description of what must be said.

Anything that is *the* explanation (in the text sense) of something is obviously an explanation of something. But everything that is *an* explanation (in the text sense) is not necessarily the explanation of anything. *An* explanation (in the text sense) of something is simply an answer to a question formed from (7) in the above way. Thus an explanation, like an answer to a question formed from (7), may be right or wrong, clear or unclear, confirmed or disconfirmed, complete or incomplete. Like an answer to such a question, its rightness or wrongness is relative not only to the question that would be cited at the place indicated by Q, but also to the known or assumed basis of a known or assumed predicament.

The remarks made earlier should be repeated here. The account of the nature of explanation just given falls short of what is eventually wanted: it fails to provide the sort of insight that can be translated into explicit standards and into a pattern of analysis applicable to all explanations and capable of deciding their correctness; it fails to make explicit the criteria that make correct explanations *correct* explanations. But as I remarked before, an adequate account that tells us all we want to know about explanations in general must wait until prior problems have been solved. This discussion merely brings us to the problems, and merely suggests an approach.