they justify. This correspondence must clearly preserve the relations of greater and less, and so make the manifold of probability relations and that of degrees of belief similar in Mr Russell's sense. I think it is a pity that Mr Keynes did not see this clearly, because the exactitude of this correspondence would have provided quite as worthy material for his scepticism as did the numerical measurement of probability relations. Indeed some of his arguments against their numerical measurement appear to apply quite equally well against their exact correspondence with degrees of belief; for instance, he argues that if rates of insurance correspond to subjective, i.e. actual, degrees of belief, these are not rationally determined, and we cannot infer that probability relations can be similarly measured. It might be argued that the true conclusion in such a case was not that, as Mr Keynes thinks, to the non-numerical probability relation corresponds a non-numerical degree of rational belief, but that degrees of belief, which were always numerical, did not correspond one to one with the probability relations justifying them. For it is, I suppose, conceivable that degrees of belief could be measured by a psychogalvanometer or some such instrument, and Mr Keynes would hardly wish it to follow that probability relations could all be derivatively measured with the measures of the beliefs which they justify.

But let us now return to a more fundamental criticism of Mr. Keynes' views, which is the obvious one that there really do not seem to be any such things as the probability relations he describes. He supposes that, at any rate in certain cases, they can be perceived; but speaking for myself I feel confident that this is not true. I do not perceive them, and if I am to be persuaded that they exist it must be by argument; moreover I shrewdly suspect that others do not perceive them either, because they are able to come to so very little agreement as to which of them relates any two given propositions.
All we appear to know about them are certain general propositions, the laws of addition and multiplication; it is as if everyone knew the laws of geometry but no one could tell whether any given object were round or square; and I find it hard to imagine how so large a body of general knowledge can be combined with so slender a stock of particular facts. It is true that about some particular cases there is agreement, but these somehow paradoxically are always immensely complicated; we all agree that the probability of a coin coming down heads is \( \frac{1}{2} \), but we can none of us say exactly what is the evidence which forms the other term for the probability relation about which we are then judging. If, on the other hand, we take the simplest possible pairs of propositions such as 'This is red' and 'That is blue' or 'This is red' and 'That is red', whose logical relations should surely be easiest to see, no one, I think, pretends to be sure what is the probability relation which connects them. Or, perhaps, they may claim to see the relation but they will not be able to say anything about it with certainty, to state if it is more or less than \( \frac{1}{2} \), or so on. They may, of course, say that it is incomparable with any numerical relation, but a relation about which so little can be truly said will be of little scientific use and it will be hard to convince a sceptic of its existence. Besides this view is really rather paradoxical; for any believer in induction must admit that between 'This is red' as conclusion and 'This is round', together with a billion propositions of the form 'a is round and red' as evidence, there is a finite probability relation; and it is hard to suppose that as we accumulate instances there is suddenly a point, say after 233 instances, at which the probability relation becomes finite and so comparable with some numerical relations.

It seems to me that if we take the two propositions 'a is red', 'b is red', we cannot really discern more than four
simple logical relations between them; namely identity of form, identity of predicate, diversity of subject, and logical independence of import. If anyone were to ask me what probability one gave to the other, I should not try to answer by contemplating the propositions and trying to discern a logical relation between them, I should, rather, try to imagine that one of them was all that I knew, and to guess what degree of confidence I should then have in the other. If I were able to do this, I might no doubt still not be content with it but might say 'This is what I should think, but, of course, I am only a fool' and proceed to consider what a wise man would think and call that the degree of probability. This kind of self-criticism I shall discuss later when developing my own theory; all that I want to remark here is that no one estimating a degree of probability simply contemplates the two propositions supposed to be related by it; he always considers inter alia his own actual or hypothetical degree of belief. This remark seems to me to be borne out by observation of my own behaviour; and to be the only way of accounting for the fact that we can all give estimates of probability in cases taken from actual life, but are quite unable to do so in the logically simplest cases in which, were probability a logical relation, it would be easiest to discern.

Another argument against Mr Keynes' theory can, I think, be drawn from his inability to adhere to it consistently even in discussing first principles. There is a passage in his chapter on the measurement of probabilities which reads as follows:—

"Probability is, vide Chapter II (§ 12), relative in a sense to the principles of human reason. The degree of probability, which it is rational for us to entertain, does not presume perfect logical insight, and is relative in part to the secondary propositions which we in fact know; and it is not dependent upon whether more perfect logical insight
is or is not conceivable. It is the degree of probability to
which those logical processes lead, of which our minds
are capable; or, in the language of Chapter II, which those
secondary propositions justify, which we in fact know. If
we do not take this view of probability, if we do not limit
it in this way and make it, to this extent, relative to human
powers, we are altogether adrift in the unknown; for we
cannot ever know what degree of probability would be justified
by the perception of logical relations which we are, and must
always be, incapable of comprehending."

This passage seems to me quite unreconcilable with the
view which Mr Keynes adopts everywhere except in this and
another similar passage. For he generally holds that the
degree of belief which we are justified in placing in the con-
clusion of an argument is determined by what relation of
probability unites that conclusion to our premisses. There
is only one such relation and consequently only one relevant
true secondary proposition, which, of course, we may or may
not know, but which is necessarily independent of the human
mind. If we do not know it, we do not know it and cannot
tell how far we ought to believe the conclusion. But often,
he supposes, we do know it; probability relations are not
ones which we are incapable of comprehending. But on this
view of the matter the passage quoted above has no mean-
ing: the relations which justify probable beliefs are pro-
bability relations, and it is nonsense to speak of them being
justified by logical relations which we are, and must always
be, incapable of comprehending.

The significance of the passage for our present purpose
lies in the fact that it seems to presuppose a different view
of probability, in which indefinable probability relations
play no part, but in which the degree of rational belief depends
on a variety of logical relations. For instance, there
might be between the premiss and conclusion the relation

1 p. 32, his italics.
that the premiss was the logical product of a thousand instances of a generalization of which the conclusion was one other instance, and this relation, which is not an indefinable probability relation but definable in terms of ordinary logic and so easily recognizable, might justify a certain degree of belief in the conclusion on the part of one who believed the premiss. We should thus have a variety of ordinary logical relations justifying the same or different degrees of belief. To say that the probability of \( a \) given \( h \) was such-and-such would mean that between \( a \) and \( h \) was some relation justifying such-and-such a degree of belief. And on this view it would be a real point that the relation in question must not be one which the human mind is incapable of comprehending.

This second view of probability as depending on logical relations but not itself a new logical relation seems to me more plausible than Mr Keynes' usual theory; but this does not mean that I feel at all inclined to agree with it. It requires the somewhat obscure idea of a logical relation justifying a degree of belief, which I should not like to accept as indefinable because it does not seem to be at all a clear or simple notion. Also it is hard to say what logical relations justify what degrees of belief, and why; any decision as to this would be arbitrary, and would lead to a logic of probability consisting of a host of so-called 'necessary' facts, like formal logic on Mr Chadwick's view of logical constants.\(^1\) Whereas I think it far better to seek an explanation of this 'necessity' after the model of the work of Mr Wittgenstein, which enables us to see clearly in what precise sense and why logical propositions are necessary, and in a general way why the system of formal logic consists of the propositions it does consist of, and what is their common characteristic. Just as natural science tries to explain and

\(^1\) J. A. Chadwick, "Logical Constants," *Mind*, 1927.
account for the facts of nature, so philosophy should try, in a sense, to explain and account for the facts of logic; a task ignored by the philosophy which dismisses these facts as being unaccountably and in an indefinable sense 'necessary'.

Here I propose to conclude this criticism of Mr Keynes' theory, not because there are not other respects in which it seems open to objection, but because I hope that what I have already said is enough to show that it is not so completely satisfactory as to render futile any attempt to treat the subject from a rather different point of view.

(3) Degrees of Belief

The subject of our inquiry is the logic of partial belief, and I do not think we can carry it far unless we have at least an approximate notion of what partial belief is, and how, if at all, it can be measured. It will not be very enlightening to be told that in such circumstances it would be rational to believe a proposition to the extent of \( \frac{2}{3} \), unless we know what sort of a belief in it that means. We must therefore try to develop a purely psychological method of measuring belief. It is not enough to measure probability; in order to apportion correctly our belief to the probability we must also be able to measure our belief.

It is a common view that belief and other psychological variables are not measurable, and if this is true our inquiry will be vain; and so will the whole theory of probability conceived as a logic of partial belief; for if the phrase 'a belief two-thirds of certainty' is meaningless, a calculus whose sole object is to enjoin such beliefs will be meaningless also. Therefore unless we are prepared to give up the whole thing as a bad job we are bound to hold that beliefs can to some extent be measured. If we were to follow the analogy
of Mr Keynes' treatment of probabilities we should say that some beliefs were measurable and some not; but this does not seem to me likely to be a correct account of the matter: I do not see how we can sharply divide beliefs into those which have a position in the numerical scale and those which have not. But I think beliefs do differ in measurability in the following two ways. First, some beliefs can be measured more accurately than others; and, secondly, the measurement of beliefs is almost certainly an ambiguous process leading to a variable answer depending on how exactly the measurement is conducted. The degree of a belief is in this respect like the time interval between two events; before Einstein it was supposed that all the ordinary ways of measuring a time interval would lead to the same result if properly performed. Einstein showed that this was not the case; and time interval can no longer be regarded as an exact notion, but must be discarded in all precise investigations. Nevertheless, time interval and the Newtonian system are sufficiently accurate for many purposes and easier to apply.

I shall try to argue later that the degree of a belief is just like a time interval; it has no precise meaning unless we specify more exactly how it is to be measured. But for many purposes we can assume that the alternative ways of measuring it lead to the same result, although this is only approximately true. The resulting discrepancies are more glaring in connection with some beliefs than with others, and these therefore appear less measurable. Both these types of deficiency in measurability, due respectively to the difficulty in getting an exact enough measurement and to an important ambiguity in the definition of the measurement process, occur also in physics and so are not difficulties peculiar to our problem; what is peculiar is that it is difficult to form any idea of how the measurement is to be conducted, how a unit is to be obtained, and so on.
Let us then consider what is implied in the measurement of beliefs. A satisfactory system must in the first place assign to any belief a magnitude or degree having a definite position in an order of magnitudes; beliefs which are of the same degree as the same belief must be of the same degree as one another, and so on. Of course this cannot be accomplished without introducing a certain amount of hypothesis or fiction. Even in physics we cannot maintain that things that are equal to the same thing are equal to one another unless we take 'equal' not as meaning 'sensibly equal' but a fictitious or hypothetical relation. I do not want to discuss the metaphysics or epistemology of this process, but merely to remark that if it is allowable in physics it is allowable in psychology also. The logical simplicity characteristic of the relations dealt with in a science is never attained by nature alone without any admixture of fiction.

But to construct such an ordered series of degrees is not the whole of our task; we have also to assign numbers to these degrees in some intelligible manner. We can of course easily explain that we denote full belief by 1, full belief in the contradictory by 0, and equal beliefs in the proposition and its contradictory by $\frac{1}{2}$. But it is not so easy to say what is meant by a belief $\frac{2}{3}$ of certainty, or a belief in the proposition being twice as strong as that in its contradictory. This is the harder part of the task, but it is absolutely necessary; for we do calculate numerical probabilities, and if they are to correspond to degrees of belief we must discover some definite way of attaching numbers to degrees of belief. In physics we often attach numbers by discovering a physical process of addition: the measure-numbers of lengths are not assigned arbitrarily subject only to the proviso that the greater length shall have the greater measure; we determine them further by deciding on a

---

physical meaning for addition; the length got by putting together two given lengths must have for its measure the sum of their measures. A system of measurement in which there is nothing corresponding to this is immediately recognized as arbitrary, for instance Mohs' scale of hardness\(^1\) in which 10 is arbitrarily assigned to diamond, the hardest known material, 9 to the next hardest, and so on. We have therefore to find a process of addition for degrees of belief, or some substitute for this which will be equally adequate to determine a numerical scale.

Such is our problem; how are we to solve it? There are, I think, two ways in which we can begin. We can, in the first place, suppose that the degree of a belief is something perceptible by its owner; for instance that beliefs differ in the intensity of a feeling by which they are accompanied, which might be called a belief-feeling or feeling of conviction, and that by the degree of belief we mean the intensity of this feeling. This view would be very inconvenient, for it is not easy to ascribe numbers to the intensities to feelings; but apart from this it seems to me observably false, for the beliefs which we hold most strongly are often accompanied by practically no feeling at all; no one feels strongly about things he takes for granted.

We are driven therefore to the second supposition that the degree of a belief is a causal property of it, which we can express vaguely as the extent to which we are prepared to act on it. This is a generalization of the well-known view, that the differentia of belief lies in its causal efficacy, which is discussed by Mr Russell in his *Analysis of Mind*. He there dismisses it for two reasons, one of which seems entirely to miss the point. He argues that in the course of trains of thought we believe many things which do not lead to action. This objection is however beside the mark, because

---

\(^1\) Ibid., p. 271.
it is not asserted that a belief is an idea which does actually lead to action, but one which would lead to action in suitable circumstances; just as a lump of arsenic is called poisonous not because it actually has killed or will kill anyone, but because it would kill anyone if he ate it. Mr Russell’s second argument is, however, more formidable. He points out that it is not possible to suppose that beliefs differ from other ideas only in their effects, for if they were otherwise identical their effects would be identical also. This is perfectly true, but it may still remain the case that the nature of the difference between the causes is entirely unknown or very vaguely known, and that what we want to talk about is the difference between the effects, which is readily observable and important.

As soon as we regard belief quantitatively, this seems to me the only view we can take of it. It could well be held that the difference between believing and not believing lies in the presence or absence of introspectible feelings. But when we seek to know what is the difference between believing more firmly and believing less firmly, we can no longer regard it as consisting in having more or less of certain observable feelings; at least I personally cannot recognize any such feelings. The difference seems to me to lie in how far we should act on these beliefs: this may depend on the degree of some feeling or feelings, but I do not know exactly what feelings and I do not see that it is indispensable that we should know. Just the same thing is found in physics; men found that a wire connecting plates of zinc and copper standing in acid deflected a magnetic needle in its neighbourhood. Accordingly as the needle was more or less deflected the wire was said to carry a larger or a smaller current. The nature of this ‘current’ could only be conjectured: what were observed and measured were simply its effects.

It will no doubt be objected that we know how strongly
TRUTH AND PROBABILITY

we believe things, and that we can only know this if we can measure our belief by introspection. This does not seem to me necessarily true; in many cases, I think, our judgment about the strength of our belief is really about how we should act in hypothetical circumstances. It will be answered that we can only tell how we should act by observing the present belief-feeling which determines how we should act; but again I doubt the cogency of the argument. It is possible that what determines how we should act determines us also directly or indirectly to have a correct opinion as to how we should act, without its ever coming into consciousness.

Suppose, however, I am wrong about this and that we can decide by introspection the nature of belief, and measure its degree; still, I shall argue, the kind of measurement of belief with which probability is concerned is not this kind but is a measurement of belief qua basis of action. This can I think be shown in two ways. First, by considering the scale of probabilities between 0 and 1, and the sort of way we use it, we shall find that it is very appropriate to the measurement of belief as a basis of action, but in no way related to the measurement of an introspected feeling. For the units in terms of which such feelings or sensations are measured are always, I think, differences which are just perceptible: there is no other way of obtaining units. But I see no ground for supposing that the interval between a belief of degree \( \frac{1}{4} \) and one of degree \( \frac{1}{3} \) consists of as many just perceptible changes as does that between one of \( \frac{1}{3} \) and one of \( \frac{4}{7} \), or that a scale based on just perceptible differences would have any simple relation to the theory of probability. On the other hand the probability of \( \frac{1}{3} \) is clearly related to the kind of belief which would lead to a bet of 2 to 1, and it will be shown below how to generalize this relation so as to apply to action in general. Secondly, the quantitative aspects of beliefs as the basis of action are evidently more important than the intensities of belief-feelings.
The latter are no doubt interesting, but may be very variable from individual to individual, and their practical interest is entirely due to their position as the hypothetical causes of beliefs *qua* bases of action.

It is possible that some one will say that the extent to which we should act on a belief in suitable circumstances is a hypothetical thing, and therefore not capable of measurement. But to say this is merely to reveal ignorance of the physical sciences which constantly deal with and measure hypothetical quantities; for instance, the electric intensity at a given point is the force which would act on a unit charge if it were placed at the point.

Let us now try to find a method of measuring beliefs as bases of possible actions. It is clear that we are concerned with dispositional rather than with actualized beliefs; that is to say, not with beliefs at the moment when we are thinking of them, but with beliefs like my belief that the earth is round, which I rarely think of, but which would guide my action in any case to which it was relevant.

The old-established way of measuring a person's belief is to propose a bet, and see what are the lowest odds which he will accept. This method I regard as fundamentally sound; but it suffers from being insufficiently general, and from being necessarily inexact. It is inexact partly because of the diminishing marginal utility of money, partly because the person may have a special eagerness or reluctance to bet, because he either enjoys or dislikes excitement or for any other reason, e.g. to make a book. The difficulty is like that of separating two different co-operating forces. Besides, the proposal of a bet may inevitably alter his state of opinion; just as we could not always measure electric intensity by actually introducing a charge and seeing what force it was subject to, because the introduction of the charge would change the distribution to be measured.
TRUTH AND PROBABILITY

In order therefore to construct a theory of quantities of belief which shall be both general and more exact, I propose to take as a basis a general psychological theory, which is now universally discarded, but nevertheless comes, I think, fairly close to the truth in the sort of cases with which we are most concerned. I mean the theory that we act in the way we think most likely to realize the objects of our desires, so that a person's actions are completely determined by his desires and opinions. This theory cannot be made adequate to all the facts, but it seems to me a useful approximation to the truth particularly in the case of our self-conscious or professional life, and it is presupposed in a great deal of our thought. It is a simple theory and one which many psychologists would obviously like to preserve by introducing unconscious desires and unconscious opinions in order to bring it more into harmony with the facts. How far such fictions can achieve the required result I do not attempt to judge: I only claim for what follows approximate truth, or truth in relation to this artificial system of psychology, which like Newtonian mechanics can, I think, still be profitably used even though it is known to be false.

It must be observed that this theory is not to be identified with the psychology of the Utilitarians, in which pleasure had a dominating position. The theory I propose to adopt is that we seek things which we want, which may be our own or other people's pleasure, or anything else whatever, and our actions are such as we think most likely to realize these goods. But this is not a precise statement, for a precise statement of the theory can only be made after we have introduced the notion of quantity of belief.

Let us call the things a person ultimately desires 'goods', and let us at first assume that they are numerically measurable and additive. That is to say that if he prefers for its own sake an hour's swimming to an hour's reading, he will prefer
two hours' swimming to one hour's swimming and one hour's reading. This is of course absurd in the given case but this may only be because swimming and reading are not ultimate goods, and because we cannot imagine a second hour's swimming precisely similar to the first, owing to fatigue, etc.

Let us begin by supposing that our subject has no doubts about anything, but certain opinions about all propositions. Then we can say that he will always choose the course of action which will lead in his opinion to the greatest sum of good.

It should be emphasized that in this essay good and bad are never to be understood in any ethical sense but simply as denoting that to which a given person feels desire and aversion.

The question then arises how we are to modify this simple system to take account of varying degrees of certainty in his beliefs. I suggest that we introduce as a law of psychology that his behaviour is governed by what is called the mathematical expectation; that is to say that, if \( p \) is a proposition about which he is doubtful, any goods or bards for whose realization \( p \) is in his view a necessary and sufficient condition enter into his calculations multiplied by the same fraction, which is called the 'degree of his belief in \( p \)’. We thus define degree of belief in a way which presupposes the use of the mathematical expectation.

We can put this in a different way. Suppose his degree of belief in \( p \) is \( \frac{m}{n} \); then his action is such as he would choose it to be if he had to repeat it exactly \( n \) times, in \( m \) of which \( p \) was true, and in the others false. [Here it may be necessary to suppose that in each of the \( n \) times he had no memory of the previous ones.]

This can also be taken as a definition of the degree of belief, and can easily be seen to be equivalent to the previous definition. Let us give an instance of the sort of case which might occur. I am at a cross-roads and do not know the way; but I rather think one of the two ways is right. I propose therefore
to go that way but keep my eyes open for someone to ask; if
now I see someone half a mile away over the fields, whether
I turn aside to ask him will depend on the relative
inconvenience of going out of my way to cross the fields or
of continuing on the wrong road if it is the wrong road. But
it will also depend on how confident I am that I am right;
and clearly the more confident I am of this the less distance
I should be willing to go from the road to check my opinion.
I propose therefore to use the distance I would be prepared
to go to ask, as a measure of the confidence of my opinion;
and what I have said above explains how this is to be done.
We can set it out as follows: suppose the disadvantage of
going \( x \) yards to ask is \( f(x) \), the advantage of arriving at the
right destination is \( r \), that of arriving at the wrong one \( w \).
Then if I should just be willing to go a distance \( d \) to ask, the
degree of my belief that I am on the right road is given by
\[
\phi = 1 - \frac{f(d)}{r - w}.
\]

For such an action is one it would just pay me to take,
if I had to act in the same way \( n \) times, in \( n\phi \) of which I was
on the right way but in the others not.

For the total good resulting from not asking each time
\[
= n\phi r + n(1 - \phi)w
\]
\[
= nw + n\phi(r - w),
\]

that resulting from asking at distance \( x \) each time
\[
= nx - nf(x). \quad [\text{I now always go right}.]
\]

This is greater than the preceding expression, provided
\[
f(x) < (r - w)(1 - \phi),
\]
\[
\therefore \text{the critical distance } d \text{ is connected with } \phi, \text{ the degree of}
\text{belief, by the relation } f(d) = (r - w)(1 - \phi)
\]
\[
or \phi = 1 - \frac{f(d)}{r - w} \quad \text{as asserted above.}
It is easy to see that this way of measuring beliefs gives results agreeing with ordinary ideas; at any rate to the extent that full belief is denoted by 1, full belief in the contradictory by 0, and equal belief in the two by \( \frac{1}{2} \). Further, it allows validity to betting as means of measuring beliefs. By proposing a bet on \( p \) we give the subject a possible course of action from which so much extra good will result to him if \( p \) is true and so much extra bad if \( p \) is false. Supposing the bet to be in goods and bads instead of in money, he will take a bet at any better odds than those corresponding to his state of belief; in fact his state of belief is measured by the odds he will just take; but this is vitiated, as already explained, by love or hatred of excitement, and by the fact that the bet is in money and not in goods and bads. Since it is universally agreed that money has a diminishing marginal utility, if money bets are to be used, it is evident that they should be for as small stakes as possible. But then again the measurement is spoiled by introducing the new factor of reluctance to bother about trifles.

Let us now discard the assumption that goods are additive and immediately measurable, and try to work out a system with as few assumptions as possible. To begin with we shall suppose, as before, that our subject has certain beliefs about everything; then he will act so that what he believes to be the total consequences of his action will be the best possible. If then we had the power of the Almighty, and could persuade our subject of our power, we could, by offering him options, discover how he placed in order of merit all possible courses of the world. In this way all possible worlds would be put in an order of value, but we should have no definite way of representing them by numbers. There would be no meaning in the assertion that the difference in value between \( a \) and \( \beta \) was equal to that between \( \gamma \) and \( \delta \). [Here and elsewhere we use Greek letters to represent the different possible totalities]
of events between which our subject chooses—the ultimate organic unities.]

Suppose next that the subject is capable of doubt; then we could test his degree of belief in different propositions by making him offers of the following kind. Would you rather have world $\alpha$ in any event; or world $\beta$ if $p$ is true, and world $\gamma$ if $p$ is false? If, then, he were certain that $p$ was true, he would simply compare $\alpha$ and $\beta$ and choose between them as if no conditions were attached; but if he were doubtful his choice would not be decided so simply. I propose to lay down axioms and definitions concerning the principles governing choices of this kind. This is, of course, a very schematic version of the situation in real life, but it is, I think, easier to consider it in this form.

There is first a difficulty which must be dealt with; the propositions like $p$ in the above case which are used as conditions in the options offered may be such that their truth or falsity is an object of desire to the subject. This will be found to complicate the problem, and we have to assume that there are propositions for which this is not the case, which we shall call ethically neutral. More precisely an atomic proposition $p$ is called ethically neutral if two possible worlds differing only in regard to the truth of $p$ are always of equal value; and a non-atomic proposition $p$ is called ethically neutral if all its atomic truth-arguments $^1$ are ethically neutral.

We begin by defining belief of degree $\frac{1}{4}$ in an ethically neutral proposition. The subject is said to have belief of degree $\frac{1}{4}$ in such a proposition $p$ if he has no preference between the options (1) $\alpha$ if $p$ is true, $\beta$ if $p$ is false, and (2) $\alpha$ if $p$ is false, $\beta$ if $p$ is true, but has a preference between $\alpha$ and $\beta$ simply. We suppose by an axiom that if this is true of any

---

$^1$ I assume here Wittgenstein's theory of propositions; it would probably be possible to give an equivalent definition in terms of any other theory.
one pair $\alpha, \beta$ it is true of all such pairs.\(^1\) This comes roughly to defining belief of degree $\frac{1}{2}$ as such a degree of belief as leads to indifference between betting one way and betting the other for the same stakes.

Belief of degree $\frac{1}{2}$ as thus defined can be used to measure values numerically in the following way. We have to explain what is meant by the difference in value between $\alpha$ and $\beta$ being equal to that between $\gamma$ and $\delta$; and we define this to mean that, if $\rho$ is an ethically neutral proposition believed to degree $\frac{1}{2}$, the subject has no preference between the options (1) $\alpha$ if $\rho$ is true, $\delta$ if $\rho$ is false, and (2) $\beta$ if $\rho$ is true, $\gamma$ if $\rho$ is false.

This definition can form the basis of a system of measuring values in the following way:—

Let us call any set of all worlds equally preferable to a given world a value: we suppose that if world $\alpha$ is preferable to $\beta$ any world with the same value as $\alpha$ is preferable to any world with the same value as $\beta$ and shall say that the value of $\alpha$ is greater than that of $\beta$. This relation 'greater than' orders values in a series. We shall use $\alpha$ henceforth both for the world and its value.

**Axioms.**

(1) There is an ethically neutral proposition $\rho$ believed to degree $\frac{1}{2}$.

(2) If $\rho, \sigma$ are such propositions and the option

$\alpha$ if $\rho$, $\delta$ if not-$\rho$ is equivalent to $\beta$ if $\rho$, $\gamma$ if not-$\rho$ then $\alpha$ if $\sigma$, $\delta$ if not-$\sigma$ is equivalent to $\beta$ if $\sigma$, $\gamma$ if not-$\sigma$.

**Def.** In the above case we say $\alpha \beta = \gamma \delta$.

**Theorems.** If $\alpha \beta = \gamma \delta$,

then $\beta \alpha = \delta \gamma$, $\alpha \gamma = \beta \delta$, $\gamma \alpha = \delta \beta$.

\(^1\) $\alpha$ and $\beta$ must be supposed so far undefined as to be compatible with both $\rho$ and not-$\rho$. 
(2a) If $a\beta = \gamma \delta$, then $a > \beta$ is equivalent to $\gamma > \delta$

and $a = \beta$ is equivalent to $\gamma = \delta$.

(3) If option $A$ is equivalent to option $B$ and $B$ to $C$ then $A$ to $C$.

Theorem. If $a\beta = \gamma \delta$ and $\beta \eta = \zeta \gamma$, then $a\eta = \zeta \delta$.

(4) If $a\beta = \gamma \delta$, $\gamma \delta = \eta \zeta$, then $a\beta = \eta \zeta$.

(5) $(a, \beta, \gamma)$. $E_1 (\forall x) (ax = \beta \gamma)$.

(6) $(a, \beta)$. $E_1 (\forall x) (ax = x\beta)$.

(7) Axiom of continuity:—Any progression has a limit (ordinal).

(8) Axiom of Archimedes.

These axioms enable the values to be correlated one-one with real numbers so that if $a^1$ corresponds to $a$, etc.

$$a\beta = \gamma \delta \equiv a^1 - \beta^1 = \gamma^1 - \delta^1.$$  

Henceforth we use $a$ for the correlated real number $a^1$ also.

Having thus defined a way of measuring value we can now derive a way of measuring belief in general. If the option of $a$ for certain is indifferent with that of $\beta$ if $\rho$ is true and $\gamma$ if $\rho$ is false,$^1$ we can define the subject's degree of belief in $\rho$ as the ratio of the difference between $a$ and $\gamma$ to that between $\beta$ and $\gamma$; which we must suppose the same for all $a$'s, $\beta$'s and $\gamma$'s that satisfy the conditions. This amounts roughly

---

$^1$ Here $\beta$ must include the truth of $\rho$, $\gamma$ its falsity; $\rho$ need no longer be ethically neutral. But we have to assume that there is a world with any assigned value in which $\rho$ is true, and one in which $\rho$ is false.
to defining the degree of belief in \( p \) by the odds at which the subject would bet on \( p \), the bet being conducted in terms of differences of value as defined. The definition only applies to partial belief and does not include certain beliefs; for belief of degree 1 in \( p \), \( a \) for certain is indifferent with \( a \) if \( p \) and any \( \beta \) if not-\( p \).

We are also able to define a very useful new idea—'the degree of belief in \( p \) given \( q \)'. This does not mean the degree of belief in 'If \( p \) then \( q \)', or that in '\( p \) entails \( q \)', or that which the subject would have in \( p \) if he knew \( q \), or that which he ought to have. It roughly expresses the odds at which he would now bet on \( p \), the bet only to be valid if \( q \) is true. Such conditional bets were often made in the eighteenth century.

The degree of belief in \( p \) given \( q \) is measured thus. Suppose the subject indifferent between the options (1) \( a \) if \( q \) true, \( \beta \) if \( q \) false, (2) \( \gamma \) if \( p \) true and \( q \) true, \( \delta \) if \( p \) false and \( q \) true, \( \beta \) if \( q \) false. Then the degree of his belief in \( p \) given \( q \) is the ratio of the difference between \( a \) and \( \delta \) to that between \( \gamma \) and \( \delta \), which we must suppose the same for any \( a, \beta, \gamma, \delta \) which satisfy the given conditions. This is not the same as the degree to which he would believe \( p \), if he believed \( q \) for certain; for knowledge of \( q \) might for psychological reasons profoundly alter his whole system of beliefs.

Each of our definitions has been accompanied by an axiom of consistency, and in so far as this is false, the notion of the corresponding degree of belief becomes invalid. This bears some analogy to the situation in regard to simultaneity discussed above.

I have not worked out the mathematical logic of this in detail, because this would, I think, be rather like working out to seven places of decimals a result only valid to two. My logic cannot be regarded as giving more than the sort of way it might work.
From these definitions and axioms it is possible to prove the fundamental laws of probable belief (degrees of belief lie between 0 and 1):

(1) Degree of belief in \( \mathcal{P} \) + degree of belief in \( \mathcal{P} \) = 1.

(2) Degree of belief in \( \mathcal{P} \) given \( q \) + degree of belief in \( \mathcal{P} \) given \( q \) = 1.

(3) Degree of belief in (\( \mathcal{P} \) and \( q \)) = degree of belief in \( \mathcal{P} \) \( \times \) degree of belief in \( q \) given \( \mathcal{P} \).

(4) Degree of belief in (\( \mathcal{P} \) and \( q \)) + degree of belief in (\( \mathcal{P} \) and \( \overline{q} \)) = degree of belief in \( \mathcal{P} \).

The first two are immediate. (3) is proved as follows.

Let degree of belief in \( \mathcal{P} = x \), that in \( q \) given \( \mathcal{P} = y \).

Then \( \xi \) for certain = \( \xi + (1 - x)t \) if \( \mathcal{P} \) true, \( \xi - xt \) if \( \mathcal{P} \) false, for any \( t \).

\[
\xi + (1 - x) \ t \text{ if } \mathcal{P} \text{ true =}
\]

\[
\begin{cases} 
\xi + (1 - x) \ t + (1 - y) \ u \text{ if '} \mathcal{P} \text{ and } q \text{' true}, \\
\xi + (1 - x) \ t - yu \text{ if } \mathcal{P} \text{ true } q \text{ false} ; & \text{for any } u.
\end{cases}
\]

Choose \( u \) so that \( \xi + (1 - x) \ t - yu = \xi - xt \),

i.e. let \( u = t/y \) (\( y \neq 0 \))

Then \( \xi \) for certain =

\[
\begin{cases} 
\xi + (1 - x) \ t + (1 - y) \ t/y \text{ if } \mathcal{P} \text{ and } q \text{ true} \\
\xi - xt \text{ otherwise},
\end{cases}
\]

\[ \therefore \text{ degree of belief in '} \mathcal{P} \text{ and } q \text{' } = \frac{xt}{t + (1 - y) \ t/y} = xy. \ (t \neq 0) \]

If \( y = 0 \), take \( t = 0 \).
Then \( \xi \) for certain \( = \xi \) if \( p \) true, \( \xi \) if \( p \) false

\[
= \xi + u \quad \text{if } p \text{ true, } q \text{ true} \; ; \; \xi \quad \text{if } p \text{ false, } q \text{ false} \; ; \; \xi \quad \text{if } p \text{ false}
\]

\[
= \xi + u, \; pq \text{ true} \; ; \; \xi, \; pq \text{ false}
\]

.: degree of belief in \( pq \) = 0.

(4) follows from (2), (3) as follows:—

Degree of belief in \( pq \) = that in \( p \times \) that in \( q \) given \( p \), by (3).
Similarly degree of belief in \( \bar{p} \bar{q} \) = that in \( p \times \) that in \( \bar{q} \) given \( p \)

.: sum = degree of belief in \( p \), by (2).

These are the laws of probability, which we have proved to be necessarily true of any consistent set of degrees of belief. Any definite set of degrees of belief which broke them would be inconsistent in the sense that it violated the laws of preference between options, such as that preferability is a transitive asymmetrical relation, and that if \( a \) is preferable to \( b \), \( b \) for certain cannot be preferable to \( a \) if \( p \), \( b \) if not-\( p \). If anyone’s mental condition violated these laws, his choice would depend on the precise form in which the options were offered him, which would be absurd. He could have a book made against him by a cunning better and would then stand to lose in any event.

We find, therefore, that a precise account of the nature of partial belief reveals that the laws of probability are laws of consistency, an extension to partial beliefs of formal logic, the logic of consistency. They do not depend for their meaning on any degree of belief in a proposition being uniquely determined as the rational one; they merely distinguish those sets of beliefs which obey them as consistent ones.

Having any definite degree of belief implies a certain measure of consistency, namely willingness to bet on a given proposition at the same odds for any stake, the stakes being measured
in terms of ultimate values. Having degrees of belief obeying the laws of probability implies a further measure of consistency, namely such a consistency between the odds acceptable on different propositions as shall prevent a book being made against you.

Some concluding remarks on this section may not be out of place. First, it is based fundamentally on betting, but this will not seem unreasonable when it is seen that all our lives we are in a sense betting. Whenever we go to the station we are betting that a train will really run, and if we had not a sufficient degree of belief in this we should decline the bet and stay at home. The options God gives us are always conditional on our guessing whether a certain proposition is true. Secondly, it is based throughout on the idea of mathematical expectation; the dissatisfaction often felt with this idea is due mainly to the inaccurate measurement of goods. Clearly mathematical expectations in terms of money are not proper guides to conduct. It should be remembered, in judging my system, that in it value is actually defined by means of mathematical expectation in the case of beliefs of degree \( \frac{1}{3} \), and so may be expected to be scaled suitably for the valid application of the mathematical expectation in the case of other degrees of belief also.

Thirdly, nothing has been said about degrees of belief when the number of alternatives is infinite. About this I have nothing useful to say, except that I doubt if the mind is capable of contemplating more than a finite number of alternatives. It can consider questions to which an infinite number of answers are possible, but in order to consider the answers it must lump them into a finite number of groups. The difficulty becomes practically relevant when discussing induction, but even then there seems to me no need to introduce it. We can discuss whether past experience gives a high probability to the sun's rising to-morrow without
bothering about what probability it gives to the sun’s rising each morning for evermore. For this reason I cannot but feel that Mr Ritchie’s discussion of the problem is unsatisfactory; it is true that we can agree that inductive generalizations need have no finite probability, but particular expectations entertained on inductive grounds undoubtedly do have a high numerical probability in the minds of all of us. We all are more certain that the sun will rise to-morrow than that I shall not throw 12 with two dice first time, i.e. we have a belief of higher degree than \( \frac{3}{8} \) in it. If induction ever needs a logical justification it is in connection with the probability of an event like this.

(4) The Logic of Consistency

We may agree that in some sense it is the business of logic to tell us what we ought to think; but the interpretation of this statement raises considerable difficulties. It may be said that we ought to think what is true, but in that sense we are told what to think by the whole of science and not merely by logic. Nor, in this sense, can any justification be found for partial belief; the ideally best thing is that we should have beliefs of degree 1 in all true propositions and beliefs of degree 0 in all false propositions. But this is too high a standard to expect of mortal men, and we must agree that some degree of doubt or even of error may be humanly speaking justified.

1 A. D. Ritchie, "Induction and Probability," Mind, 1926, p. 318. 'The conclusion of the foregoing discussion may be simply put. If the problem of induction be stated to be "How can inductive generalizations acquire a large numerical probability?" then this is a pseudo-problem, because the answer is "They cannot". This answer is not, however, a denial of the validity of induction but is a direct consequence of the nature of probability. It still leaves untouched the real problem of induction which is "How can the probability of an induction be increased?" and it leaves standing the whole of Keynes' discussion on this point.'
TRUTH AND PROBABILITY

Many logicians, I suppose, would accept as an account of their science the opening words of Mr Keynes' *Treatise on Probability*: "Part of our knowledge we obtain direct; and part by argument. The Theory of Probability is concerned with that part which we obtain by argument, and it treats of the different degrees in which the results so obtained are conclusive or inconclusive." Where Mr Keynes says 'the Theory of Probability', others would say Logic. It is held, that is to say, that our opinions can be divided into those we hold immediately as a result of perception or memory, and those which we derive from the former by argument. It is the business of Logic to accept the former class and criticize merely the derivation of the second class from them.

Logic as the science of argument and inference is traditionally and rightly divided into deductive and inductive; but the difference and relation between these two divisions of the subject can be conceived in extremely different ways. According to Mr Keynes valid deductive and inductive arguments are fundamentally alike; both are justified by logical relations between premiss and conclusion which differ only in degree. This position, as I have already explained, I cannot accept. I do not see what these inconclusive logical relations can be or how they can justify partial beliefs. In the case of conclusive logical arguments I can accept the account of their validity which has been given by many authorities, and can be found substantially the same in Kant, De Morgan, Peirce and Wittgenstein. All these authors agree that the conclusion of a formally valid argument is contained in its premisses; that to deny the conclusion while accepting the premisses would be self-contradictory; that a formal deduction does not increase our knowledge, but only brings out clearly what we already know in another form; and that we are bound to accept its validity on pain of being
inconsistent with ourselves. The logical relation which justifies the inference is that the sense or import of the conclusion is contained in that of the premisses.

But in the case of an inductive argument this does not happen in the least; it is impossible to represent it as resembling a deductive argument and merely weaker in degree; it is absurd to say that the sense of the conclusion is partially contained in that of the premisses. We could accept the premisses and utterly reject the conclusion without any sort of inconsistency or contradiction.

It seems to me, therefore, that we can divide arguments into two radically different kinds, which we can distinguish in the words of Peirce as (1) 'explicative, analytic, or deductive' and (2) 'ampliative, synthetic, or (loosely speaking) inductive'.1 Arguments of the second type are from an important point of view much closer to memories and perceptions than to deductive arguments. We can regard perception, memory and induction as the three fundamental ways of acquiring knowledge; deduction on the other hand is merely a method of arranging our knowledge and eliminating inconsistencies or contradictions.

Logic must then fall very definitely into two parts: (excluding analytic logic, the theory of terms and propositions) we have the lesser logic, which is the logic of consistency, or formal logic; and the larger logic, which is the logic of discovery, or inductive logic.

What we have now to observe is that this distinction in no way coincides with the distinction between certain and partial beliefs; we have seen that there is a theory of consistency in partial beliefs just as much as of consistency in certain beliefs, although for various reasons the former is not so important as the latter. The theory of probability is in fact a generalization of formal logic; but in the process

1 C. S. Peirce, Chance Love and Logic, p. 92.
of generalization one of the most important aspects of formal logic is destroyed. If $p$ and $q$ are inconsistent so that $q$ follows logically from $p$, that $p$ implies $q$ is what is called by Wittgenstein a 'tautology' and can be regarded as a degenerate case of a true proposition not involving the idea of consistency. This enables us to regard (not altogether correctly) formal logic including mathematics as an objective science consisting of objectively necessary propositions. It thus gives us not merely the $\text{ἀνάγκη λέγεω}$, that if we assert $p$ we are bound in consistency to assert $q$ also, but also the $\text{ἀνάγκη ἐλναύ}$, that if $p$ is true, so must $q$ be. But when we extend formal logic to include partial beliefs this direct objective interpretation is lost; if we believe $pq$ to the extent of $\frac{1}{3}$, and $p\bar{q}$ to the extent of $\frac{1}{3}$, we are bound in consistency to believe $\bar{p}$ also to the extent of $\frac{1}{8}$. This is the $\text{ἀνάγκη λέγεω}$; but we cannot say that if $pq$ is $\frac{1}{3}$ true and $p\bar{q}$ $\frac{1}{3}$ true, $\bar{p}$ also must be $\frac{1}{8}$ true, for such a statement would be sheer nonsense. There is no corresponding $\text{ἀνάγκη ἐλναύ}$. Hence, unlike the calculus of consistent full belief, the calculus of objective partial belief cannot be immediately interpreted as a body of objective tautology.

This is, however, possible in a roundabout way; we saw at the beginning of this essay that the calculus of probabilities could be interpreted in terms of class-ratios; we have now found that it can also be interpreted as a calculus of consistent partial belief. It is natural, therefore, that we should expect some intimate connection between these two interpretations, some explanation of the possibility of applying the same mathematical calculus to two such different sets of phenomena. Nor is an explanation difficult to find; there are many connections between partial beliefs and frequencies. For instance, experienced frequencies often lead to corresponding partial beliefs, and partial beliefs lead to the expectation of corresponding frequencies in accordance with Bernouilli's
Theorem. But neither of these is exactly the connection we want; a partial belief cannot in general be connected uniquely with any actual frequency, for the connection is always made by taking the proposition in question as an instance of a propositional function. What propositional function we choose is to some extent arbitrary and the corresponding frequency will vary considerably with our choice. The pretensions of some exponents of the frequency theory that partial belief means full belief in a frequency proposition cannot be sustained. But we found that the very idea of partial belief involves reference to a hypothetical or ideal frequency; supposing goods to be additive, belief of degree \( \frac{m}{n} \) is the sort of belief which leads to the action which would be best if repeated \( n \) times in \( m \) of which the proposition is true; or we can say more briefly that it is the kind of belief most appropriate to a number of hypothetical occasions otherwise identical in a proportion \( \frac{m}{n} \) of which the proposition in question is true. It is this connection between partial belief and frequency which enables us to use the calculus of frequencies as a calculus of consistent partial belief. And in a sense we may say that the two interpretations are the objective and subjective aspects of the same inner meaning, just as formal logic can be interpreted objectively as a body of tautology and subjectively as the laws of consistent thought.

We shall, I think, find that this view of the calculus of probability removes various difficulties that have hitherto been found perplexing. In the first place it gives us a clear justification for the axioms of the calculus, which on such a system as Mr Keynes' is entirely wanting. For now it is easily seen that if partial beliefs are consistent they will obey these axioms, but it is utterly obscure why Mr Keynes'
mysterious logical relations should obey them.\footnote{It appears in Mr Keynes' system as if the principal axioms—the laws of addition and multiplication—were nothing but definitions. This is merely a logical mistake; his definitions are formally invalid unless corresponding axioms are presupposed. Thus his definition of multiplication presupposes the law that if the probability of a given \( bh \) is equal to that of \( c \) given \( dh \), and the probability of \( b \) given \( h \) is equal to that of \( d \) given \( h \), then will the probabilities of \( ab \) given \( h \) and of \( cd \) given \( h \) be equal.} We should be so curiously ignorant of the instances of these relations, and so curiously knowledgeable about their general laws.

Secondly, the Principle of Indifference can now be altogether dispensed with; we do not regard it as belonging to formal logic to say what should be a man's expectation of drawing a white or a black ball from an urn; his original expectations may within the limits of consistency be any he likes; all we have to point out is that if he has certain expectations he is bound in consistency to have certain others. This is simply bringing probability into line with ordinary formal logic, which does not criticize premisses but merely declares that certain conclusions are the only ones consistent with them. To be able to turn 'the Principle of Indifference out of formal logic is a great advantage; for it is fairly clearly impossible to lay down purely logical conditions for its validity, as is attempted by Mr Keynes. I do not want to discuss this question in detail, because it leads to hair-splitting and arbitrary distinctions which could be discussed for ever. But anyone who tries to decide by Mr Keynes' methods what are the proper alternatives to regard as equally probable in molecular mechanics, e.g. in Gibbs' phase-space, will soon be convinced that it is a matter of physics rather than pure logic. By using the multiplication formula, as it is used in inverse probability, we can on Mr Keynes' theory reduce all probabilities to quotients of \textit{a priori} probabilities; it is therefore in regard to these latter that the Principle of Indifference is of primary importance; but here the question is obviously not one of formal logic. How can we on merely
logical grounds divide the spectrum into equally probable bands?

A third difficulty which is removed by our theory is the one which is presented to Mr Keynes' theory by the following case. I think I perceive or remember something but am not sure; this would seem to give me some ground for believing it, contrary to Mr Keynes' theory, by which the degree of belief in it which it would be rational for me to have is that given by the probability relation between the proposition in question and the things I know for certain. He cannot justify a probable belief founded not on argument but on direct inspection. In our view there would be nothing contrary to formal logic in such a belief; whether it would be reasonable would depend on what I have called the larger logic which will be the subject of the next section; we shall there see that there is no objection to such a possibility, with which Mr Keynes' method of justifying probable belief solely by relation to certain knowledge is quite unable to cope.

(5) The Logic of Truth

The validity of the distinction between the logic of consistency and the logic of truth has been often disputed; it has been contended on the one hand that logical consistency is only a kind of factual consistency; that if a belief in \( p \) is inconsistent with one in \( q \), that simply means that \( p \) and \( q \) are not both true, and that this is a necessary or logical fact. I believe myself that this difficulty can be met by Wittgenstein's theory of tautology, according to which if a belief in \( p \) is inconsistent with one in \( q \), that \( p \) and \( q \) are not both true is not a fact but a tautology. But I do not propose to discuss this question further here.

From the other side it is contended that formal logic or the logic of consistency is the whole of logic, and inductive
logic either nonsense or part of natural science. This contention, which would I suppose be made by Wittgenstein, I feel more difficulty in meeting. But I think it would be a pity, out of deference to authority, to give up trying to say anything useful about induction.

Let us therefore go back to the general conception of logic as the science of rational thought. We found that the most generally accepted parts of logic, namely, formal logic, mathematics and the calculus of probabilities, are all concerned simply to ensure that our beliefs are not self-contradictory. We put before ourselves the standard of consistency and construct these elaborate rules to ensure its observance. But this is obviously not enough; we want our beliefs to be consistent not merely with one another but also with the facts\(^1\): nor is it even clear that consistency is always advantageous; it may well be better to be sometimes right than never right. Nor when we wish to be consistent are we always able to be: there are mathematical propositions whose truth or falsity cannot as yet be decided. Yet it may humanly speaking be right to entertain a certain degree of belief in them on inductive or other grounds: a logic which proposes to justify such a degree of belief must be prepared actually to go against formal logic; for to a formal truth formal logic can only assign a belief of degree 1. We could prove in Mr Keynes' system that its probability is 1 on any evidence. This point seems to me to show particularly clearly that human logic or the logic of truth, which tells men how they should think, is not merely independent of but sometimes actually incompatible with formal logic.

In spite of this nearly all philosophical thought about human logic and especially induction has tried to reduce it in some way

to formal logic. Not that it is supposed, except by a very few, that consistency will of itself lead to truth; but consistency combined with observation and memory is frequently credited with this power.

Since an observation changes (in degree at least) my opinion about the fact observed, some of my degrees of belief after the observation are necessarily inconsistent with those I had before. We have therefore to explain how exactly the observation should modify my degrees of belief; obviously if \( p \) is the fact observed, my degree of belief in \( q \) after the observation should be equal to my degree of belief in \( q \) given \( p \) before, or by the multiplication law to the quotient of my degree of belief in \( pq \) by my degree of belief in \( p \). When my degrees of belief change in this way we can say that they have been changed consistently by my observation.

By using this definition, or on Mr Keynes' system simply by using the multiplication law, we can take my present degrees of belief, and by considering the totality of my observations, discover from what initial degrees of belief my present ones would have arisen by this process of consistent change. My present degrees of belief can then be considered logically justified if the corresponding initial degrees of belief are logically justified. But to ask what initial degrees of belief are justified, or in Mr Keynes' system what are the absolutely \textit{a priori} probabilities, seems to me a meaningless question; and even if it had a meaning I do not see how it could be answered.

If we actually applied this process to a human being, found out, that is to say, on what \textit{a priori} probabilities his present opinions could be based, we should obviously find them to be ones determined by natural selection, with a general tendency to give a higher probability to the simpler alternatives. But, as I say, I cannot see what could be meant by
asking whether these degrees of belief were logically justified. Obviously the best thing would be to know for certain in advance what was true and what false, and therefore if any one system of initial beliefs is to receive the philosopher's approbation it should be this one. But clearly this would not be accepted by thinkers of the school I am criticising. Another alternative is to apportion initial probabilities on the purely formal system expounded by Wittgenstein, but as this gives no justification for induction it cannot give us the human logic which we are looking for.

Let us therefore try to get an idea of a human logic which shall not attempt to be reducible to formal logic. Logic, we may agree, is concerned not with what men actually believe, but what they ought to believe, or what it would be reasonable to believe. What then, we must ask, is meant by saying that it is reasonable for a man to have such and such a degree of belief in a proposition? Let us consider possible alternatives.

First, it sometimes means something explicable in terms of formal logic: this possibility for reasons already explained we may dismiss. Secondly, it sometimes means simply that were I in his place (and not e.g. drunk) I should have such a degree of belief. Thirdly, it sometimes means that if his mind worked according to certain rules, which we may roughly call 'scientific method', he would have such a degree of belief. But fourthly it need mean none of these things; for men have not always believed in scientific method, and just as we ask 'But am I necessarily reasonable', we can also ask 'But is the scientist necessarily reasonable?' In this ultimate meaning it seems to me that we can identify reasonable opinion with the opinion of an ideal person in similar circumstances. What, however, would this ideal person's opinion be? As has previously been remarked, the highest ideal would be always to have a true
opinion and be certain of it; but this ideal is more suited to God than to man.¹

We have therefore to consider the human mind and what is the most we can ask of it.³ The human mind works essentially according to general rules or habits; a process of thought not proceeding according to some rule would simply be a random sequence of ideas; whenever we infer A from B we do so in virtue of some relation between them. We can therefore state the problem of the ideal as "What habits in a general sense would it be best for the human mind to have?" This is a large and vague question which could hardly be answered unless the possibilities were first limited by a fairly definite conception of human nature. We could imagine some very useful habits unlike those possessed by any men. [It must be explained that I use habit in the most general possible sense to mean simply rule or law of behaviour, including instinct: I do not wish to distinguish acquired

¹ [Earlier draft of matter of preceding paragraph in some ways better.—F.P.R.

What is meant by saying that a degree of belief is reasonable? First and often that it is what I should entertain if I had the opinions of the person in question at the time but was otherwise as I am now, e.g. not drunk. But sometimes we go beyond this and ask: 'Am I reasonable?' This may mean, do I conform to certain enumerable standards which we call scientific method, and which we value on account of those who practise them and the success they achieve. In this sense to be reasonable means to think like a scientist, or to be guided only by ratiocination and induction or something of the sort (i.e. reasonable means reflective). Thirdly, we may go to the root of why we admire the scientist and criticize not primarily an individual opinion but a mental habit as being conducive or otherwise to the discovery of truth or to entertaining such degrees of belief as will be most useful. (To include habits of doubt or partial belief.) Then we can criticize an opinion according to the habit which produced it. This is clearly right because it all depends on this habit; it would not be reasonable to get the right conclusion to a syllogism by remembering vaguely that you leave out a term which is common to both premisses.

We use reasonable in sense 1 when we say of an argument of a scientist this does not seem to me reasonable; in sense 2 when we contrast reason and superstition or instinct; in sense 3 when we estimate the value of new methods of thought such as soothsaying.]

³ What follows to the end of the section is almost entirely based on the writings of C. S. Peirce. [Especially his "Illustrations of the Logic of Science", Popular Science Monthly, 1877 and 1878, reprinted in Chance, Love and Logic (1923).]
rules or habits in the narrow sense from innate rules or instincts, but propose to call them all habits alike.] A completely general criticism of the human mind is therefore bound to be vague and futile, but something useful can be said if we limit the subject in the following way.

Let us take a habit of forming opinion in a certain way; e.g. the habit of proceeding from the opinion that a toadstool is yellow to the opinion that it is unwholesome. Then we can accept the fact that the person has a habit of this sort, and ask merely what degree of opinion that the toadstool is unwholesome it would be best for him to entertain when he sees it; i.e. granting that he is going to think always in the same way about all yellow toadstools, we can ask what degree of confidence it would be best for him to have that they are unwholesome. And the answer is that it will in general be best for his degree of belief that a yellow toadstool is unwholesome to be equal to the proportion of yellow toadstools which are in fact unwholesome. (This follows from the meaning of degree of belief.) This conclusion is necessarily vague in regard to the spatio-temporal range of toadstools which it includes, but hardly vaguer than the question which it answers. (Cf. density at a point of gas composed of molecules.)

Let us put it in another way: whenever I make an inference, I do so according to some rule or habit. An inference is not completely given when we are given the premiss and conclusion; we require also to be given the relation between them in virtue of which the inference is made. The mind works by general laws; therefore if it infers \( q \) from \( \phi \), this will generally be because \( q \) is an instance of a function \( \phi \) and \( \phi \) the corresponding instance of a function \( \psi \) such that the mind would always infer \( \phi \) from \( \psi \). When therefore we criticize not opinions but the processes by which they are formed, the rule of the inference determines for us a range to which the frequency theory can be applied. The rule of the inference
may be narrow, as when seeing lightning I expect thunder, or wide, as when considering 99 instances of a generalization which I have observed to be true I conclude that the 100th is true also. In the first case the habit which determines the process is ‘After lightning expect thunder’; the degree of expectation which it would be best for this habit to produce is equal to the proportion of cases of lightning which are actually followed by thunder. In the second case the habit is the more general one of inferring from 99 observed instances of a certain sort of generalization that the 100th instance is true also; the degree of belief it would be best for this habit to produce is equal to the proportion of all cases of 99 instances of a generalization being true, in which the 100th is true also.

Thus given a single opinion, we can only praise or blame it on the ground of truth or falsity: given a habit of a certain form, we can praise or blame it accordingly as the degree of belief it produces is near or far from the actual proportion in which the habit leads to truth. We can then praise or blame opinions derivatively from our praise or blame of the habits that produce them.

This account can be applied not only to habits of inference but also to habits of observation and memory; when we have a certain feeling in connection with an image we think the image represents something which actually happened to us, but we may not be sure about it; the degree of direct confidence in our memory varies. If we ask what is the best degree of confidence to place in a certain specific memory feeling, the answer must depend on how often when that feeling occurs the event whose image it attaches to has actually taken place.

Among the habits of the human mind a position of peculiar importance is occupied by induction. Since the time of Hume a great deal has been written about the justification for inductive inference. Hume showed that it could not
be reduced to deductive inference or justified by formal logic. So far as it goes his demonstration seems to me final; and the suggestion of Mr Keynes that it can be got round by regarding induction as a form of probable inference cannot in my view be maintained. But to suppose that the situation which results from this is a scandal to philosophy is, I think, a mistake.

We are all convinced by inductive arguments, and our conviction is reasonable because the world is so constituted that inductive arguments lead on the whole to true opinions. We are not, therefore, able to help trusting induction, nor if we could help it do we see any reason why we should, because we believe it to be a reliable process. It is true that if any one has not the habit of induction, we cannot prove to him that he is wrong; but there is nothing peculiar in that. If a man doubts his memory or his perception we cannot prove to him that they are trustworthy; to ask for such a thing to be proved is to cry for the moon, and the same is true of induction. It is one of the ultimate sources of knowledge just as memory is: no one regards it as a scandal to philosophy that there is no proof that the world did not begin two minutes ago and that all our memories are not illusory.

We all agree that a man who did not make inductions would be unreasonable: the question is only what this means. In my view it does not mean that the man would in any way sin against formal logic or formal probability; but that he had not got a very useful habit, without which he would be very much worse off, in the sense of being much less likely¹ to have true opinions.

This is a kind of pragmatism: we judge mental habits by whether they work, i.e. whether the opinions they lead

¹ 'Likely' here simply means that I am not sure of this, but only have a certain degree of belief in it.
to are for the most part true, or more often true than those which alternative habits would lead to.

Induction is such a useful habit, and so to adopt it is reasonable. All that philosophy can do is to analyse it, determine the degree of its utility, and find on what characteristics of nature this depends. An indispensable means for investigating these problems is induction itself, without which we should be helpless. In this circle lies nothing vicious. It is only through memory that we can determine the degree of accuracy of memory; for if we make experiments to determine this effect, they will be useless unless we remember them.

Let us consider in the light of the preceding discussion what sort of subject is inductive or human logic—the logic of truth. Its business is to consider methods of thought, and discover what degree of confidence should be placed in them, i.e. in what proportion of cases they lead to truth. In this investigation it can only be distinguished from the natural sciences by the greater generality of its problems. It has to consider the relative validity of different types of scientific procedure, such as the search for a causal law by Mill's Methods, and the modern mathematical methods like the *a priori* arguments used in discovering the Theory of Relativity. The proper plan of such a subject is to be found in Mill\(^1\); I do not mean the details of his Methods or even his use of the Law of Causality. But his way of treating the subject as a body of inductions about inductions, the Law of Causality governing lesser laws and being itself proved by induction by simple enumeration. The different scientific methods that can be used are in the last resort judged by induction by simple enumeration; we choose the simplest law that fits the facts, but unless we found that laws so obtained also fitted facts other than those they were made to fit, we should discard this procedure for some other.

---

\(^1\) Cf. also the account of 'general rules' in the Chapter 'Of Unphilosophical Probability' in Hume's *Treatise*.
VI

FURTHER CONSIDERATIONS (1928)

A. REASONABLE DEGREE OF BELIEF

When we pass beyond reasonable = my, or = scientific, to define it precisely is quite impossible. Following Peirce we predicate it of a habit not of an individual judgment. Roughly, reasonable degree of belief = proportion of cases in which habit leads to truth. But in trying to be more exact we encounter the following difficulties:

(1) We cannot always take the actual habit: this may be correctly derived from some previous accidentally misleading experience. We then look to wider habit of forming such a habit.

(2) We cannot take proportion of actual cases; e.g. in a card game very rarely played, so that of the particular combination in question there are very few actual instances.

(3) We sometimes really assume a theory of the world with laws and chances, and mean not the proportion of actual cases but what is chance on our theory.

(4) But it might be argued that this complication was not necessary on account of (1) by which we only consider very general habits of which there are so many instances that, if chance on our theory differed from the actual proportion, our theory would have to be wrong.

(5) Also in an ultimate case like induction, there could be no chance for it: it is not the sort of thing that has a chance.

Fortunately there is no point in fixing on a precise sense of ‘reasonable’; this could only be required for one of two
reasons: either because the reasonable was the subject-matter of a science (which is not the case); or because it helped us to be reasonable to know what reasonableness is (which it does not, though some false notions might hinder us). To make clear that it is not needed for either of these purposes we must consider (1) the content of logic

and (2) the utility of logic.

THE CONTENT OF LOGIC

(1) Preliminary philosophico-psychological investigation into nature of thought, truth and reasonableness.

(2) Formulae for formal inference = mathematics.

(3) Hints for avoiding confusion (belongs to medical psychology).

(4) Outline of most general propositions known or used as habits of inference from an abstract point of view; either crudely inductive, as 'Mathematical method has solved all these other problems, therefore . . . ', or else systematic, when it is called metaphysics. All this might anyhow be called metaphysics; but it is regarded as logic when adduced as bearing on an unsolved problem, not simply as information interesting for its own sake.

The only one of these which is a distinct science is evidently (2).

THE UTILITY OF LOGIC

That of (1) above and of (3) are evident: the interesting ones are (2) and (4). (2) = mathematics is indispensable for manipulating and systematizing our knowledge. Besides this (2) and (4) help us in some way in coming to conclusions in judgment.
FURTHER CONSIDERATIONS

LOGIC AS SELF-CONTROL (Cf. Peirce)

Self-control in general means either

(1) not acting on the temporarily uppermost desire, but stopping to think it out; i.e. pay regard to all desires and see which is really stronger; its value is to eliminate inconsistency in action;

or (2) forming as a result of a decision habits of acting not in response to temporary desire or stimulus but in a definite way adjusted to permanent desire.

The difference is that in (1) we stop to think it out but in (2) we’ve thought it out before and only stop to do what we had previously decided to do.

So also logic enables us

(1) Not to form a judgment on the evidence immediately before us, but to stop and think of all else that we know in any way relevant. It enables us not to be inconsistent, and also to pay regard to very general facts, e.g. all crows I’ve seen are black, so this one will be—No; colour is in such and such other species a variable quality. Also e.g. not merely to argue from $\phi a . \phi b \ldots$ to $(x) . \phi x$ probable, but to consider the bearing of $a, b \ldots$ are the class I’ve seen (and visible ones are specially likely or unlikely to be $\phi$). This difference between biassed and random selection.\(^1\)

(2) To form certain fixed habits of procedure or interpretation only revised at intervals when we think things out. In this it is the same as any general judgment; we should only regard the process as 'logic' when it is very general, not e.g. to expect a woman to be unfaithful, but e.g. to disregard correlation coefficients with a probable error greater than themselves.

With regard to forming a judgment or a partial judgment

\(^1\) Vide infra 'Chance'.
(which is a decision to have a belief of such a degree, i.e. to act in a certain way) we must note:—

(a) What we ask is ‘$\rho$?’ not ‘Would it be true to think $\rho$?’ nor ‘Would it be reasonable to think $\rho$?’ (But these might be useful first steps.)

but (b) ‘Would it be true to think $\rho$?’ can never be settled without settling $\rho$ to which it is equivalent.

(c) ‘Would it be reasonable to think $\rho$?’ means simply ‘Is $\rho$ what usually happens in such a case?’ and is as vague as ‘usually’. To put this question may help us, but it will often seem no easier to answer than $\rho$ itself.

(d) Nor can the precise sense in which ‘reasonable’ or ‘usually’ can usefully be taken be laid down, nor weight assigned on any principle to different considerations of such a sort. E.g. the death-rate for men of 60 is $\frac{1}{10}$, but all the 20 red-haired 60-year-old men I’ve known have lived till 70. What should I expect of a new red-haired man of 60? I can but put the evidence before me, and let it act on my mind. There is a conflict of two ‘usually’s’ which must work itself out in my mind; one is not the really reasonable, the other the really unreasonable.

(e) When, however, the evidence is very complicated, statistics are introduced to simplify it. They are to be chosen in such a way as to influence me as nearly as possible in the same way as would the whole facts they represent if I could apprehend them clearly. But this cannot altogether be reduced to a formula; the rest of my knowledge may affect the matter; thus $\rho$ may be equivalent in influence to $q$, but not $\rho h$ to $q h$.

(f) There are exceptional cases in which ‘It would be reasonable to think $\rho$’ absolutely settles the matter. Thus if we are told that one of these people's names begins with A and that there are 8 of them, it is reasonable to believe to degree $\frac{1}{8}$th that any particular one's name begins with A,
and this is what we should all do (unless we felt there was something else relevant).

(g) Nevertheless, to introduce the idea of ‘reasonable’ is really a mistake; it is better to say ‘usually’, which makes clear the vagueness of the range: what is reasonable depends on what is taken as relevant; if we take enough as relevant, whether it is reasonable to think $p$ becomes at least as difficult a question as $p$. If we take everything as relevant, they are the same.

(h) What ought we to take as relevant? Those sorts of things which it is useful to take as relevant; if we could rely on being reasonable in regard to what we do take as relevant, this would mean everything. Otherwise it is impossible to say; but the question is one asked by a spectator not by the thinker himself: if the thinker feels a thing relevant he can’t dismiss it; and if he feels it irrelevant he can’t use it.

(i) Only then if we in fact feel very little to be relevant, do or can we answer the question by an appeal to what is reasonable, this being then equivalent to what we know and consider relevant.

(j) What are or are not taken as relevant are not only propositions but formal facts, e.g. $a = a$: we may react differently to $\phi a$ than to any other $\phi x$ not because of anything we know about $a$ but e.g. for emotional reasons.
B. STATISTICS

The science of statistics is concerned with abbreviating facts about numerous individuals which are interpreted as a random selection from an infinite 'population'. If the qualities concerned are discrete, this means simply that we consider the proportions of the observed individuals which have the qualities, and ascribe these proportions to the hypothetical population. If the qualities are continuous, we take the population to be of a convenient simple form containing various parameters which are then chosen to give the highest probability to the instances observed. In either case the probable error is calculated for such a sample from such a population. (For all this see Fisher.) ¹

The significance of this procedure is that we record in a convenient simple form

(1) The approximate proportions having the given qualities in different degrees,

(2) The number of instances which we have observed (the weight of our induction) (probable error).

For the use of the figures to give a degree of belief with regard to a new instance no rule can be given.

The introduction of an infinite population is a stupid fiction, which cannot be defended except by some reference to proceeding to a limit, which destroys its sense. The procedure of calculating parameters by maximum likelihood and probable error can be defined as a process in pure mathematics; its significance is in suggesting a theory or set

of chances. Proportion of infinite population should be replaced by chance.

Of course the purpose is not always simple induction but causal analysis: we find the chances are not what we expect, therefore the die is biassed or people are more careful nowadays etc.
C. CHANCE

(1) There are no such things as objective chances in the sense in which some people imagine there are, e.g. N. Campbell, Nisbet.¹

There is, for instance, no established fact of the form ‘In \( n \) consecutive throws the number of heads lies between \( \frac{n}{2} \pm \varepsilon(n) \)’. On the contrary we have good reason to believe that any such law would be broken if we took enough instances of it.

Nor is there any fact established empirically about infinite series of throws; this formulation is only adopted to avoid contradiction by experience; and what no experience can contradict, none can confirm, let alone establish.

(N. Campbell makes a simple mistake about this.)

A crude frequency theory is ruled out because it justifies the ‘maturity of odds’ argument, e.g. in regard to sex of offspring.

(2) Hence chances must be defined by degrees of belief; but they do not correspond to anyone’s actual degrees of belief; the chances of 1,000 heads, and of 999 heads followed by a tail, are equal, but everyone expects the former more than the latter.

(3) Chances are degrees of belief within a certain system of beliefs and degrees of belief; not those of any actual person, but in a simplified system to which those of actual people, especially the speaker, in part approximate.

(4) This system of beliefs consists, firstly, of natural laws,


206
which are in it believed for certain, although, of course, people are not really quite certain of them.

(5) Besides these the system contains various things of this sort: when knowing $\psi x$ and nothing else relevant, always expect $\phi x$ with degree of belief $\phi$ (what is or is not relevant is also specified in the system); which is also written the chance of $\phi$ given $\psi$ is $\phi$ (if $\phi = 1$ it is the same as a law). These chances together with the laws form a deductive system according to the rules of probability, and the actual beliefs of a user of the system should approximate to those deduced from a combination of the system and the particular knowledge of fact possessed by the user, this last being (inexactly) taken as certain.

(6) The chances in such a system must not be confounded with frequencies; the chance of $\phi x$ given $\psi x$ might be different even from the known frequency of $\psi$'s which are $\phi$'s. E.g. the chance of a coin falling heads yesterday is $\frac{1}{2}$ since 'yesterday' is irrelevant, but the proportion that actually fell heads yesterday might be 1.

(7) It is, however, obvious that we are not armed with systems giving us a degree of belief in every possible proposition for any basis of factual knowledge. Our systems only cover part of the field; and where we have no system we say we do not know the chances.

(8) The phenomena for which we have systematic chances are games of chance, births, deaths, and all sorts of correlation coefficients.

(9) What we mean by objective chance is not merely our having in our system a chance $\frac{\phi(x)}{\psi(x)}$, but our having no hope of modifying our system into a pair of laws $ax . \psi x . \exists_x . \phi x : \beta x . \psi x . \exists_x . \sim \phi x$, etc., where $ax, \beta x$ are disjunctions of readily observable properties (previous in time to $\phi x$). This
occurs, as Poincaré points out,\(^1\) when small causes produce large effects.

Chances are in another sense objective, in that everyone agrees about them, as opposed e.g. to odds on horses.

(10) What we mean by an event not being a coincidence, or not being due to chance, is that if we came to know it, it would make us no longer regard our system as satisfactory, although on our system the event may be no more improbable than any alternative. Thus 1,000 heads running would not be due to chance; i.e. if we observed it we should change our system of chances for that penny. If it is called \(h\), the chances in our system with \(h\) as hypothesis are markedly different from our actual degrees of belief in things given \(h\).

By saying a thing is not due to chance, we only mean that our system of chances must be changed, not that it must become a system of laws. Thus for a biassed coin to come down heads is not due to chance even though it doesn’t always do so; e.g. chance may = \(\frac{3}{4}\) say, not \(\frac{1}{4}\).

If we say ‘Our meeting was not due to chance’, i.e. \textit{designed}, design is simply a factor modifying the chances; it might also be e.g. that we walk in the same road.

(11) This is why N. Campbell thinks coincidences cannot be allowed to occur; i.e. coincidences \(\mathcal{D}\) system wrong, \(\therefore\) system \(\mathcal{C}\) no coincidences. Apparently formally conclusive; but this is a mistake because the system is not a proposition which is true or false, but an imperfect approximation to a state of mind where imperfections can in certain circumstances become particularly glaring.

(12) By things being \textit{ultimately} due to chance, we mean that there is no law (here generalization of no more than manageable complexity), known or unknown, which determines the future from the past. If we suppose further that they have ultimate

\(^1\) See \textit{Science et Hypothèse} and \textit{Science et Méthode}.
chances, this means a sort of best possible system in which they have these chances.

(13) In choosing a system we have to compromise between two principles: subject always to the proviso that the system must not contradict any facts we know, we choose (other things being equal) the simplest system, and (other things being equal) we choose the system which gives the highest chance to the facts we have observed. This last is Fisher's 'Principle of Maximum Likelihood', and gives the only method of verifying a system of chances.

(14) Probability in Physics means chance as here explained, with possibly some added complexity because we are concerned with a 'theory' in Campbell's sense, not merely an ordinary system which is a generalization of Campbell's 'law.' What chance in a theory is can hardly be explained until we know more about the nature of theories.¹

(15) Statistical science must be briefly dealt with from our point of view; it has three parts:

(a) Collection and arrangement of selections from multitudinous data.

(b) Induction = forming a system of chances from the data by means of the Principle of Maximum Likelihood.

(c) Causal analysis; e.g. this die falls so often this way up, therefore its centre of gravity must be displaced towards the opposite face.

(16) The only difficulty presented is in connection with (c) causal analysis, in which we seem to take a statement of chances as a fact, and to argue 'Its falling so often six is not due to chance ' ∴ chance > $\frac{1}{6}$ ' ∴ c.g. displaced'. Reasoning which seems incompatible with our solution of the paradox that chance = $\frac{1}{6}$ is inconsistent with this coincidence, which was that 'chance = $\frac{1}{6}$ ', 'chance > $\frac{1}{6}$ ', were not propositions

¹ [See next section.—Ed.]
and so could not serve as premisses or conclusions of arguments.

(17) The difficulty is removed by the reflection that the system we are ultimately using not only gives us degree of belief or chance of $x$ falling six given $x$ is tossed $= \frac{1}{6}$, but also that that of $x$ falling six given $x$ is tossed and biassed $> \frac{1}{6}$. Consequently by transposition $x$ is biassed / $x$ falls six. $x$ is tossed $> x$ is biassed / $x$ is tossed. If $a/bh > a/h$, then $b/ah > b/h$, and this is how we are arguing. The chance of a $x$ falling six is $p$ seems to be treated as a genuine proposition, but what is really meant is an unexpressed condition, which on our system when added to the hypothesis makes the chance $p$.

(18) We can state it this way: statistical causal analysis presupposes a fundamental system within which it moves and which it leaves unchanged; this neither is nor appears to be treated like a proposition. What appears to be so treated is a narrower system derived or derivable from the fundamental system by the addition of an empirical premiss, and what is really treated as a proposition and modified or rejected is not the narrower system but the empirical premiss on which it is based.

Of course this empirical premiss may be unknown or very vaguely known; e.g. I conclude from the fact that more boys are born than girls to some superiority in the number, mobility or capacity for fertilization of male-bearing spermatazoa or one of a thousand other possible causes, because by the Principle of Indifference, which is part of my fundamental system, the observed inequality would be so unlikely if there were no such difference. But there seems no fundamental difference between this case and the biassed coin.

(19) Note on Poincaré's problem 'Why are chance events subject to law?' The fundamental answer to this is that they are not, taking the whole field of chance events no generalizations about them are possible (consider e.g. infectious diseases,
dactyIs in hexameters, deaths from horse kicks, births of great men).

Poincaré says it is paradoxical that the actuary can from ignorance derive so easily such useful conclusions whereas if he knew the laws of health he would have to go through endless calculations. In fact it is not from ignorance that he works, but from experience of frequencies.

(20) Note on 'random'.

Keynes\(^1\) gives a substantially correct account of this. But

(a) It is essential to bring in the notion of a description. What we want is not \(a\) is a random member of \(\mathcal{A}(Sx)\) for the purpose of \(\phi x\), but the description \((\forall x)(\psi x)\) is a random description when \(x = (\forall x)(\psi x)\) is irrelevant to \(\phi x/Sx \cdot h\).

(b) It is essential to extend the term to cover not merely a selection of one term but of many; thus, \(\psi \alpha\) gives a random selection of \(n\) S's with regard to \(\phi \alpha\) means \(a = \alpha(\psi x)\) is irrelevant to probabilities of the form: Proportion of \(a\) which is \(\phi = \lambda/a\in . a\in \mathcal{A}(Sx) \cdot h\).

The idea of random selection is useful in induction, where the value of the argument 'A proportion \(\lambda\) of \(\psi\)S's are \(\phi\)'s.' . 'A proportion \(\lambda\) of S's are \(\phi\)'s' depends on whether \(\psi\) is a random selector. If \(\lambda = 1\) of course the value of the argument is strengthened if \(\psi\) is biased against \(\phi\), weakened if \(\psi\) is biased in favour of it.

\(^1\) Treatise on Probability, p. 291.
A. THEORIES

Let us try to describe a theory simply as a language for discussing the facts the theory is said to explain. This need not commit us on the philosophical question of whether a theory is only a language, but rather if we knew what sort of language it would be if it were one at all, we might be further towards discovering if it is one. We must try to make our account as general as possible, but we cannot be sure that we have in fact reached the most general type of theory, since the possible complication is infinite.

First, let us consider the facts to be explained. These occur in a universe of discourse which we will call the primary system, this system being composed of all the terms and propositions (true or false) in the universe in question. We must suppose the primary system in some way given to us so that we have a notation capable of expressing every proposition in it. Of what sort must this notation be?

It might in the first case consist of names of different types any two or more of which conjoined together gave an atomic proposition; for instance, the names $a, b \ldots z$, ‘red’, ‘before’. But I think the systems we try to explain are rarely of this kind; if for instance we are concerned with a series of experiences, we do not try to explain their time order (which we could not explain by anything simpler) or

---

1 The ‘universe’ of the primary system might contain ‘blue or red’ but not ‘blue’ or ‘red’; i.e. we might be out to explain when a thing was ‘blue or red’ as opposed to ‘green or yellow’, but not which it was, blue or red. ‘Blue-or-red’ would then be a term: ‘blue’, ‘red’ nonsense for our present purpose.
even, assuming an order, whether it is \( a \) or \( b \) that comes first; we take for granted that they are in order and that \( a \) comes before \( b \), etc., and try to explain which is red, which blue, etc. \( a \) is essentially one before \( b \), and \( 'a', 'b' \), etc., are not really names but descriptions except in the case of the present. We take it for granted that these descriptions describe uniquely, and instead of \( 'a was red' \) we have e.g. \( 'The 3rd one ago was red' \). The symbols we want are not names but numbers: the 0th (i.e. the present), 1st, -1th, etc., in general the \( n \)th, and we can use red \( n \) to mean the \( n \)th is red counting forward or backward from a particular place. If the series terminates at say 100, we could write \( N(101) \), and generally \( N(m) \) if \( m > 100 \), meaning \( 'There is no \( m \)th' \); or else simply regard e.g. red \( m \) as nonsense if \( m > 100 \), whereas if we wrote \( N(m) \) we should say red \( m \) was false. I am not sure this is necessary, but it seems to me always so in practice; i.e. the terms of our primary system have a structure, and any structure can be represented by numbers (or pairs or other combinations of numbers).

It may be possible to go further than this, for of the terms in our primary system not merely some but even all may be best symbolized by numbers. For instance, colours have a structure, in which any given colour may be assigned a place by three numbers, and so on. Even smells may be so treated: the presence of the smell being denoted by 1, the absence by 0 (or all total smell qualities may be given numbers). Of course, we cannot make a proposition out of numbers without some link. Moment 3 has colour 1 and smell 2 must be written \( \chi(3) = 1 \) and \( \phi(3) = 2 \), \( \chi \) and \( \phi \) corresponding to the general forms of colour and smell, and possibly being functions with a limited number of values, so that e.g. \( \phi(3) = 55 \) might be nonsense, since there was no 55th smell.

Whether or no this is possible, it is not so advantageous where we have relatively few terms (e.g. a few smells) to
deal with. Where we have a multitude as e.g. with times, we cannot name them, and our theory will not explain a primary system in which they have names, for it will take no account of their individuality but only of their position. In general nothing is gained and clarity may be lost by using numbers when the order, etc., of the numbers corresponds to nothing in the nature of the terms.

If all terms were represented by numbers, the propositions of the primary system would all take the form of assertions about the values taken by certain one-valued numerical functions. These would not be mathematical functions in the ordinary sense; for that such a function had such-and-such a value would always be a matter of fact, not a matter of mathematics.

We have spoken as if the numbers involved were always integers, and if the finitists are right this must indeed be so in the ultimate primary system, though the integers may, of course, take the form of rationals. This means we may be concerned with pairs \((m, n)\) with \((\lambda m, \lambda n)\) always identical to \((m, n)\). If, however, our primary system is already a secondary system from some other theory, real numbers may well occur.

So much for the primary system; now for the theoretical construction.

We will begin by taking a typical form of theory, and consider later whether or not this form is the most general. Suppose the atomic propositions of our primary system are such as \(A(n), B(m, n)\) . . . where \(m, n,\) etc., take positive or negative integral values subject to any restrictions, e.g. that in \(B(m, n)\) \(m\) may only take the values 1, 2.

Then we introduce new propositional functions \(a(n), \beta(n), \gamma(m, n),\) etc., and by propositions of the secondary system we shall mean any truth-functions of the values of \(a, \beta, \gamma,\) etc. We shall also lay down propositions about these values, e.g. \((n). \ a(n). \ \beta(n)\) which we shall call axioms, and whatever
propositions of the secondary system can be deduced from the axioms we shall call theorems.

Besides this we shall make a dictionary which takes the form of a series of definitions of the functions of the primary system \( A, B, C \ldots \) in terms of those of the secondary system \( \alpha, \beta, \gamma \), e.g. \( A(n) = \alpha(n) \cdot \nu \cdot \gamma(0, n^2) \). By taking these 'definitions' as equivalences and adding them to the axioms we may be able to deduce propositions in the primary system which we shall call laws if they are general propositions, consequences if they are singular. The totality of laws and consequences will be the eliminant when \( \alpha, \beta, \gamma \ldots \), etc., are eliminated from the dictionary and axioms, and it is this totality of laws and consequences which our theory asserts to be true.

We may make this clearer by an example\(^1\); let us interpret numbers \( n, n_1, n_2 \), etc., as instants of time and suppose the primary system to contain the following functions:

\[
A(n) = \text{I see blue at } n.
\]

\[
B(n) = \text{I see red at } n.
\]

\[
[A(n) \cdot \Bar{B}(n) = \text{I see nothing at } n].
\]

\[
C(n) = \text{Between } n-1 \text{ and } n \text{ I feel my eyes open.}
\]

\[
D(n) = \text{Between } n-1 \text{ and } n \text{ I feel my eyes shut.}
\]

\[
E(n) = \text{I move forward a step at } n.
\]

\[
F(n) = \text{I move backward a step at } n.
\]

and that we construct a theory in the following way:

\(^1\) [The example seems futile, therefore try to invent a better; but it in fact brings out several good points, which it would be difficult otherwise to bring out. It may however miss some points which we will consider later. A defect in all Nicod's examples is that they do not give an external world in which anything happens.—F. P. R.]
First $m$ will be understood to take only the values 1, 2, 3,

\[ \begin{align*}
&f(1) = 2 \\
&f(2) = 3 \\
&f(3) = 1
\end{align*} \]

and $f(m)$ is defined by

Then we introduce

\[ \alpha(n, m) = \text{At time } n \text{ I am at place } m. \]

\[ \beta(n, m) = \text{At time } n \text{ place } m \text{ is blue.} \]

\[ \gamma(n) = \text{At time } n \text{ my eyes are open.} \]

And the axioms

\[ (n, m, m') : \quad \alpha(n, m) \cdot \alpha(n, m') \cdot \mathfrak{C} \cdot m = m'. \]

\[ (n). \quad (\exists m). \quad \alpha(n, m). \]

\[ (n). \quad \beta(n, 1). \]

\[ (n) : \quad \beta(n, 2) \cdot \gamma(n) \cdot \beta(n + 1, 2). \]

And the dictionary

\[ A(n) = (\exists m) \cdot \alpha(n, m) \cdot \beta(n, m) \cdot \gamma(n). \]

\[ B(n) = (\exists m) \cdot \alpha(n, m) \cdot \beta(n, m) \cdot \gamma(n). \]

\[ C(n) = \gamma(n - 1) \cdot \gamma(n). \]

\[ D(n) = \gamma(n - 1) \cdot \gamma(n). \]

\[ E(n) = (\exists m) \cdot \alpha(n - 1, m) \cdot \alpha(n, f(m)) \]

\[ F(n) = (\exists m) \cdot \alpha(n - 1, f(m)) \cdot \alpha(n, m). \]

This theory can be said to represent me as moving among 3 places, 'forwards' being in the sense $ABCA$, 'backwards' $ACBA$. Place $A$ is always blue, place $B$ alternately blue and red, place $C$ blue or red according to a law I have not discovered. If my eyes are open I see the colour of the place
I am in, if they are shut I see no colour. The laws resulting from the theory can be expressed as follows:

(1) \((n) \cdot \{\bar{A}(n) \lor \bar{B}(n)\} \cdot \{\bar{C}(n) \lor \bar{D}(n)\} \cdot \{\bar{E}(n) \lor \bar{F}(n)\}\)

(2) \((n_1, n_2) \cdot \{n_1 > n_2 \cdot C(n_1) \cdot C(n_2) \cdot \exists \cdot (\forall n_3) \cdot n_1 > n_3 > n_2 \cdot D(n_3)\}\)

\(2^1\) (2) with the C’s and D’s interchanged.

Let us define \((n_1, n_2)\) to mean

\[
\begin{array}{l}
1 \\
2
\end{array}
\]

\[\bar{N}c' \cdot \nu \{n_1 < \nu < n_2 \cdot E(\nu)\}\]

\[-\bar{N}c' \cdot \nu \{n_1 < \nu < n_2 \cdot F(\nu)\} = 0 \text{ (mod 3)}\]

\[
\begin{array}{l}
1 \\
2
\end{array}
\]

(3) \(([\exists n_1) \cdot C(n_1) \cdot n_1 < n \cdot n > \nu > n_1 \cdot \exists \cdot \bar{D}(\nu)] \text{ } \exists_n : \\
A(n) \lor \bar{B}(n)\]

(4) \(([\exists n_1) \cdot D(n_1) \cdot n_1 < n \cdot n > \nu > n_1 \cdot \exists \cdot \bar{C}(\nu)] \text{ } \exists_n : \\
\bar{A}(n) \lor \bar{B}(n)\]

(5) \((n) : (\exists m) : m = 0, 1, \text{ or } 2 : m(\nu, n) \lor \exists \bar{A}(\nu, n) : (m + 1) (\nu_1, n) . (m + 1) (\nu_2, n) \cdot \nu_1 \not= \nu_2 \text{ (mod 2)}\)

\[\lor \nu_1 . \nu_2 . \bar{A}(\nu_1) \lor \bar{A}(\nu_2) \lor \bar{B}(\nu_1) \lor \bar{B}(\nu_2).\]

[Where \(2 + 1 = 0\) for this purpose.]

These can then be compared with the axioms and dictionary, and there is no doubt that to the normal mind the axioms and dictionary give the laws in a more manageable form.

Let us now put it all into mathematics by writing

\[A(n) \quad \text{as } \phi(n) = 1\]

\[B(n) \quad \text{as } \phi(n) = -1\]
\[ \tilde{A}(n) \cdot \tilde{B}(n) \text{ as } \phi(n) = 0 \]
\[ C(n) \quad \text{as } \chi(n) = 1 \]
\[ D(n) \quad \text{as } \chi(n) = -1 \]
\[ \tilde{C}(n) \cdot \tilde{D}(n) \text{ as } \chi(n) = 0 \]
\[ E(n) \quad \text{as } \psi(n) = 1 \]
\[ F(n) \quad \text{as } \psi(n) = -1 \]
\[ \tilde{E}(n) \tilde{F}(n) \quad \text{as } \psi(n) = 0. \]

Instead of \( a(n, m) \) have \( a(n) \) a function taking values 1, 2, 3

\[ \beta(n, m) \quad ,, \quad \beta(n, m) \quad ,, \quad 1, -1 \]
\[ \gamma(n) \quad \gamma(n) \quad 1, 0 \]

Our axioms are just

1. \( (n) \cdot a(n) = 1 \lor 2 \lor 3 \)
2. \( (n) \cdot \beta(n, 1) = 1 \)
3. \( (n) \cdot \beta(n, 2) \neq \beta(n + 1, 2) \)
4. \( (n, m) \cdot \beta(n, m) = 1 \lor -1 \)
5. \( (n) \cdot \gamma(n) = 0 \lor 1 \)

Of these (1) (4) (5) hardly count since they merely say what values the functions are capable of taking.

Our definitions become.

(i) \( \phi(n) = \gamma(n) \times \beta(n, a(n)) \)

(ii) \( \chi(n) = \gamma(n) - \gamma(n - 1) \)

(iii) \( \psi(n) = \text{Remainder mod 3 of } a(n) - a(n - 1) \)

Our laws are of course that \( \phi, \chi, \psi \) must be such that \( \alpha, \beta, \gamma \) can be found to satisfy 1-5, i–iii. Going through the old laws we have instead of

1. \( \phi(n) = -1 \lor 0 \lor 1, \chi(n) = -1 \lor 0 \lor 1, \psi(n) = -1 \lor 0 \lor 1 \)

[understood].
(2) \((n, m) \cdot \sum_{r = n}^{m} \chi(r) \leq 1.\)

(3) \((\exists m) \cdot \sum_{r = m}^{n} \chi(r) = 1 : \mathcal{D}_{n} \cdot \phi(n) \neq 0.\)

(4) \((\exists m) \cdot \sum_{r = m}^{n} \chi(r) = -1 : \mathcal{D}_{n} \cdot \phi(n) = 0.\)

(5) \((n) \cdot (\exists m) \cdot \sum_{r = n'}^{n''} \psi(r) \equiv m \pmod{3} \cdot \mathcal{D}_{n'} \cdot \phi(n') \neq -1 \cdot \mathcal{D}_{n'} \cdot \phi(n') \phi(n'') = 0 \lor -1.\)

So far we have only shown the genesis of laws; consequences arise when we add to the axioms a proposition involving e.g. a particular value of \(n\), from which we can deduce propositions in the primary system not of the form \((n) \ldots \). These we call the consequences.

If we take it in its mathematical form we can explain the idea of a theory as follows: Instead of saying simply what we know about the values of the functions with which we are concerned, we say that they can be constructed in a definite way given by the dictionary out of functions satisfying certain conditions given by the axioms.

Such then is an example of a theory; before we go on to discuss systematically the different features of the example and whether they occur in any theory, let us take some questions that might be asked about theories and see how they would be answered in the present case.

1. Can we say anything in the language of this theory that we could not say without it?

Obviously not; for we can easily eliminate the functions of the second system and so say in the primary system all that the theory gives us.
2. Can we reproduce the structure of our theory by means of explicit definitions within the primary system?

[This question is important because Russell, Whitehead, Nicod and Carnap all seem to suppose that we can and must do this.]

Here there are some distinctions to make. We might, for instance, argue as follows. Supposing the laws and consequences to be true, the facts of the primary system must be such as to allow functions to be defined with all the properties of those of the secondary system, and these give the solution of our problem. But the trouble is that the laws and consequences can be made true by a number of different sets of facts, corresponding to each of which we might have different definitions. So that our problem of finding a single set of definitions which will make the dictionary and axioms true whenever the laws and consequences are true, is still unsolved. We can, however, at once solve it formally, by disjoining the sets of definitions previously obtained; i.e. if the different sets of facts satisfying the laws and consequences are \( P_1, P_2, P_3 \); and the corresponding definitions of \( a(n, m) \) are

\[
\begin{align*}
a(n, m) &= L_1 \{A, B, C \ldots, n, m\} \\
L_2 \{A, B, C, \ldots, n, m\} & \text{ etc.}
\end{align*}
\]

we make the definition

\[
\begin{align*}
a(n, m) &= P_1 \supseteq L_1 \{A, B, C \ldots n, m\}. \\
P_2 \supseteq L_2 \{A, B, C \ldots n, m\}.
\end{align*}
\]

etc.

Such a definition is formally valid and evidently fulfils our requirements.

\(^1\) Jean Nicod, *La Géométrie dans le Monde Sensible* (1924), translated in his *Problems of Geometry and Induction* (1930); Rudolf Carnap, *Der Logische Aufbau der Welt* (1928).
What can be objected to it is complexity and arbitrariness, since \( L_1, L_2 \ldots \) can probably be chosen each in many ways.

Also it explicitly assumes that our primary system is finite and contains a definite number of assignable atomic propositions.

Let us therefore see what other ways there are of proceeding.

We might at first sight suppose that the key lay simply in the dictionary; this gives definitions of \( A, B, C \ldots \) in terms of \( \alpha, \beta, \gamma \ldots \). Can we not invert it to get definitions of \( \alpha, \beta, \gamma \ldots \) in terms of \( A, B, C \ldots \)? Or, in the mathematical form, can we not solve the equations for \( \alpha, \beta, \gamma \ldots \) in terms of \( \phi, \chi, \psi \ldots \), at any rate if we add to the dictionary, as we legitimately can, those laws and axioms which merely state what values the functions are capable of taking?

When, however, we look at these equations (i), (ii), (iii) what we find is this: If we neglect the limitations on the values of the functions they possess an integral solution provided \( \gamma(n) \) can be found from (ii) so as always to be a factor of \( \phi(n) \), i.e. in general always to be \( \pm 1 \) or \( 0 \) and never to vanish unless \( \phi(n) \) vanishes. This is, of course, only true in virtue of the conditions laid on \( \phi \) and \( \chi \) by the laws; assuming these laws and the limitation on values, we get the solution

\[
\alpha(n) = \sum_{o}^{n} \psi(n) + C_1 \pmod{3}
\]

\[
\gamma(n) = \sum_{o}^{n} \chi(n) + C_2
\]

for \( \alpha \) and \( \gamma \).

And for \( \beta(n, m) \) no definite solution but e.g. the trivial one \( \beta(n, m) = \phi(n) \) (assuming \( \gamma(n) = 1 \) or \( 0 \)).

Here \( C_2 \) must be chosen so as to make \( \gamma(n) \) always \( 1 \) or \( 0 \); and the value necessary for this purpose depends on the facts.
of the primary system and cannot be deduced simply from the laws. It must in fact be one or nought:

(a) If there is a least positive or zero $n$ for which $\chi(n) \neq 0$, according as $\chi(n)$ for that $n$ is $-1$ or $+1$.

(b) If there is a least negative $n$ for which $\chi(n) \neq 0$, according as $\chi(n)$ for that $n$ is $+1$ or $-1$.

(c) If for no $n$ $\chi(n) \neq 0$ it does not matter whether $C_2$ is $+1$ or $-1$.

We thus have a disjunctive definition of $C_2$ and so of $\gamma(n)$. Again although any value of $C_1$ will satisfy the limitations on the value of $\alpha(n)$, probably only one such will satisfy the axioms, and this value will again have to be disjunctively defined. And, thirdly, $\beta(n, m)$ is not at all fixed by the equations, and it will be a complicated matter in which we shall again have to distinguish cases, to say which of the many possible solutions for $\beta(n, m)$ will satisfy the axioms.

We conclude, therefore, that there is neither in this case nor in general any simple way of inverting the dictionary so as to get either a unique or an obviously preeminent solution which will also satisfy the axioms, the reason for this lying partly in difficulties of detail in the solution of the equations, partly in the fact that the secondary system has a higher multiplicity, i.e. more degrees of freedom, than the primary. In our case the primary system contains three one valued functions, the secondary virtually five [$\beta(n, 1)$, $\beta(n, 2)$, $\beta(n, 3)$, $\alpha(n)$, $\gamma(n)$] each taking 2 or 3 values, and such an increase of multiplicity is, I think, a universal characteristic of useful theories.

Since, therefore, the dictionary alone does not suffice, the next hopeful method is to use both dictionary and axioms in a way which is referred to in many popular discussions of theories when it is said that the meaning of a proposition about the external world is what we should ordinarily regard
as the criterion or test of its truth. This suggests that we should define propositions in the secondary system by their criteria in the primary.

In following this method we have first to distinguish the sufficient criterion of a proposition from its necessary criterion. If \( \phi \) is a proposition of the secondary system, we shall mean by its sufficient criterion, \( \sigma(\phi) \), the disjunction of all propositions \( q \) of the primary system such that \( \phi \) is a logical consequence of \( q \) together with the dictionary and axioms, and such that \( \sim q \) is not a consequence of the dictionary and axioms.\(^1\) On the other hand, by the necessary criterion of \( \phi \), \( \tau(\phi) \) we shall mean the conjunction of all those propositions of the primary system which follow from \( \phi \) together with the dictionary and axioms.

We can elucidate the connection of \( \sigma(\phi) \) and \( \tau(\phi) \) as follows. Consider all truth-possibilities of atomic propositions in the primary system which are compatible with the dictionary and axioms. Denote such a truth-possibility by \( r \), the dictionary and axioms by \( a \). Then \( \sigma(\phi) \) is the disjunction of every \( r \) such that

\[
\tau(\phi) \text{ the disjunction of every } r \text{ such that } \quad r \not\models a \text{ is not a contradiction.}
\]

If we denote by \( L \) the totality of laws and consequences, i.e. the disjunction of every \( r \) here in question, then we have evidently

\[
\sigma(\phi) : \equiv : L \to \sim \tau(\sim \phi),
\]

1. The laws and consequences need not be added, since they follow from the dictionary and axioms. It might be thought, however, that we should take them instead of the axioms, but it is easy to see that this would merely increase the divergence between sufficient and necessary criteria and in general the difficulties of the method. The last clause could be put as that \( \sim q \) must not follow from or be a law or consequence.
\[ \tau(p) : = : L \cdot \sim \sigma(\sim p) \]  
\[ \sigma(p) \vee \tau(\sim p) \cdot = \cdot L. \]  

We have also

\[ \sigma(p_1 \cdot p_2) : = : \sigma(p_1) \cdot \sigma(p_2), \]  

for \( p_1 \cdot p_2 \) follows from \( q \) when and only when \( p_1 \) and \( p_2 \) both follow.

Whence, or similarly, we get the dual

\[ \tau(p_1 \vee p_2) \cdot = \cdot \tau(p_1) \vee \tau(\sim p_2). \]  

We also have

\[ \sigma(p) \supset \tau(p), \]  

(Consider the \( r \)'s above.)

\[ \sigma(p) \vee \sigma(\sim p) \cdot \supset \cdot L \cdot \supset \cdot \tau(p) \vee \tau(\sim p), \]  

(from iii)

and from (vi), (ii), (iii).

\[ \sigma(p) \cdot \supset \cdot \sim \sigma(\sim p) \cdot L, \]  

(viii)

\[ L \cdot \sim \tau(\sim p) \cdot C \cdot \tau(p). \]  

(ix)

Lastly we have

\[ \sigma(p_1) \vee \sigma(p_2) \cdot C \cdot \sigma(p_1 \vee p_2). \]  

(x)

Since if \( q \) follows either from \( p_1 \) or from \( p_2 \) it follows from \( p_1 \vee p_2 \); and the dual

\[ \tau(p_1 \cdot p_2) \cdot C \cdot \tau(p_1) \cdot \tau(p_2). \]  

(xi)

On the other hand, and this is a very important point, the converses of (vi)-(xi) are not in general true. Let us illustrate this by taking (x) and considering this ' \( r \) :

\[ B(0) \cdot \bar{A}(0) : n \neq 0 \cdot D_n \cdot \bar{A}(n) \cdot \bar{B}(n). \]
\( C(n) \equiv n \cdot n = 0 : D(n) \equiv n \cdot n = 1. \)

\( (n) \cdot \mathcal{E}n \cdot \mathcal{F}n, \)

i.e. that the man's eyes are only open once when he sees blue.

From this we can deduce \( a(0, 2) \lor a(0, 3) \)

\[ \therefore \text{ This } r \supset \sigma \{a(0, 2) \lor a(0, 3)\}. \]

But we cannot deduce from it \( a(0, 2) \) or \( a(0, 3) \), since it is equally compatible with either. Hence neither \( \sigma \{a(0, 2)\} \) nor \( \sigma \{a(0, 3)\} \) is true. Hence we do not have

\[ \sigma \{a(0, 2) \lor a(0, 3)\} \supset \sigma \{a(0, 2)\} \lor \sigma \{a(0, 3)\}. \]

It follows from this that we cannot give definitions such that, if \( p \) is any proposition of the secondary system, \( p \) will in virtue of the definitions mean \( \sigma(p) \) [or alternatively \( \tau(p) \)], for if \( p_1 \) is defined to mean \( \sigma(p_1) \), \( p_2 \) to mean \( \sigma(p_2) \), \( p_1 \lor p_2 \) will mean \( \sigma(p_1) \lor \sigma(p_2) \), which is not, in general, the same as \( \sigma(p_1 \lor p_2) \). We can therefore only use \( \sigma \) to define some of the propositions of the secondary systems, what we might call **atomic** secondary propositions, from which the meanings of the others would follow.

For instance, taking our functions \( a, \beta, \gamma \) we could proceed as follows:

\( \gamma(n) \) is defined as \( A(n) \lor B(n) \), where there is no difficulty as \( A(n) \lor B(n) = \sigma(\gamma(n)) = \tau(\gamma(n)) \).

\( \beta(n, m) \) could be defined to mean \( \sigma(\beta(n, m)) \), i.e. we should say place \( m \) was 'blue' at time \( n \), only if there were proof that it was. Otherwise we should say it was not 'blue' ('red' in common parlance).

\( \tilde{\beta}(n, m) \) would then mean \( \sigma(\tilde{\beta}(n, m)) \) not \( \sigma(\tilde{\beta}(n, m)) \).

Alternatively we could use \( \tau \), and define

\( \beta(n, m) \) to be \( \tau \{\beta(n, m)\} \).

and \( \tilde{\beta}(n, m) \) would be \( \tilde{\tau} \{\beta(n, m)\} \).
In this case we should say \( m \) was 'blue' whenever there was no proof that it was not; this could, however, have been achieved by means of \( \sigma \) if we had defined \( \beta(n, m) \) to be \( \sim \beta'(n, m) \), and \( \beta'(n, m) \) to be \( \sigma(\beta'(n, m)) \), i.e. applied \( \sigma \) to \( \beta' \) instead of \( \beta \).

In general it is clear that \( \tau \) always gives what could be got from applying \( \sigma \) to the contradictory, and we may confine our attention to \( \sigma \).

It makes, however, a real difference whether we define \( \beta \) or \( \tilde{\beta} \) by means of \( \sigma \), especially in connection with place 3. For we have no law as to the values of \( \beta(n, 3) \), nor any way of deducing one except when \( a(n, 3) \) is true and \( A(n) \) or \( B(n) \) is true.

If we define \( \beta(n, 3) \) to be \( \sigma(\beta(n, 3)) \), we shall say that 3 is never blue except when we observe it to be; if we define \( \tilde{\beta}(n, 3) \) to be \( \sigma(\tilde{\beta}(n, 3)) \) we shall say it is always blue except when we observe it not to be.

Coming now to \( a(n, m) \) we could define

\[
a(n, 1) = \sigma(a(n, 1)),
\]

\[
a(n, 2) = \sigma(a(n, 1) \lor a(n, 2)) \cdot \tilde{\sigma}(a(n, 1)),
\]

\[
a(n, 3) = \tilde{\sigma}(a(n, 1) \lor a(n, 2));
\]

and we should for any \( n \) have one and only one of \( a(n, 1) \), \( a(n, 2) \), \( a(n, 3) \) true: whereas if we simply put

\[
a(n, m) = \sigma(a(n, m)),
\]

this would not follow, since

\[
\sigma(a(n, 1)), \sigma(a(n, 2)), \sigma(a(n, 3))
\]

could quite well all be false.

\[
[E.g. \text{if } (n) \cdot \tilde{A}(n) \cdot \tilde{B}(n)]
\]

Of course in all these definitions we must suppose \( \sigma(a(n, m)) \) etc., replaced by what on calculation we find them to be. As
they stand the definitions look circular, but are not when interpreted in this way.

For instance \( \sigma(a(n, 1)) \) is \( L \), i.e. laws \((1)-(5)\) together with

\[
\exists n_1, n_2 \cdot 2(n_1, n) \cdot 2(n_2, n) \cdot n_1 \not\equiv n_2 \pmod{2} \cdot Bn_1 \cdot Bn_2.
\]

v. \( \exists n_1, n_2, n_3 \cdot 2(n_1, n) \cdot 2(n_2, n) \cdot 2(n_3, n) \cdot n_1 \not\equiv n_2 \pmod{2} \).

\[
A_{n_1} \cdot A_{n_2} \cdot B_{n_3}.
\]

v. \( \exists n_1, n_2 \cdot 1(n_1, n) \cdot 2(n_2, n) \cdot Bn_1 \cdot Bn_2 \)

Such then seem to be the definitions to which we are led by the popular phrase that the meaning of a statement in the second system is given by its criterion in the first. Are they such as we require?

What we want is that, using these definitions, the axioms and dictionary should be true whenever the theory is applicable, i.e. whenever the laws and consequences are true; i.e. that interpreted by means of these definitions, the axioms and dictionary should follow from the laws and consequences.

It is easy to see that they do not so follow. Take for instance the last axiom on p. 216:

\[
(n) : \beta(n, 2) \cdot = \cdot \beta(n + 1, 2),
\]

which means according to our definitions

\[
(n) : \sigma(\beta(n, 2)) \cdot = \cdot \sigma(\beta(n + 1, 2)),
\]

which is plainly false, since if, as is perfectly possible, the man has never opened his eyes at place 2, both \( \sigma(\beta(n, 2)) \) and \( \sigma(\beta(n + 1, 2)) \) will be false.

[The definition by \( \tau \) is no better, since \( \tau(\beta(n, 2)) \) and \( \tau(\beta(n + 1, 2)) \) would both be true.]

This line of argument is, however, exposed to an objection of the following sort: If we adopt these definitions it is true that the axioms will not follow from the laws and consequences, but it is not really necessary that they should. For the
laws and consequences cannot represent the whole empirical (i.e. primary system) basis of the theory. It is, for instance, compatible with the laws and consequences that the man should never have had his eyes open at place 2; but how could he then have ever formulated this theory with the peculiar law of alternation which he ascribes to place 2? What we want in order to construct our theory by means of explicit definitions is not that the axioms should follow from the laws and consequences alone, but from them together with certain existential propositions of the primary system representing experiences the man must have had in order to be able with any show of reason to formulate the theory.

Reasonable though this objection is in the present case, it can be seen by taking a slightly more complicated theory to provide us with no general solution of the difficulty; that is to say, such propositions as could in this way be added to the laws and consequences would not always provide a sufficient basis for the axioms. For instance, suppose the theory provided for a whole system of places identified by the movement sequences necessary to get from one to another, and it was found and embodied in the theory that the colour of each place followed a complicated cycle, the same for each place, but that the places differed from one another as to the phase of this cycle according to no ascertainable law. Clearly such a theory could be reasonably formed by a man who had not had his eyes open at each place, and had no grounds for thinking that he ever would open his eyes at all the places or even visit them at all. Suppose then $m$ is a place he never goes to, and that $\beta(n, m)$ is a function of the second system, signifying that $m$ is blue at $n$; then unless he knows the phase of $m$, we can never have $\sigma(\beta(n, m))$, but if e.g. the cycle gives a blue colour once in six, we must have from an axiom $\beta(0, m) \lor \beta(1, m) \lor \ldots \lor \beta(6, m)$. We have, therefore, just the same difficulty as before.
If, therefore, our theory is to be constructed by explicit definitions, these cannot be simple definitions by means of \( \sigma \) (or \( \tau \)), but must be more complicated. For instance, in regard to place 2 in our original example we can define

\[
\beta(0, 2) \text{ as } \sigma(\beta(0, 2)),
\]

\[
\beta(n, 2) \text{ as } \sigma(\beta(0, 2)) \text{ if } n \text{ is even},
\]

\[
\sim \sigma(\beta(0, 2)) \text{ if } n \text{ is odd}.
\]

I.e. if we do not know which phase it is, we assume it to be a certain one, including that 'assumption' in our definition. E.g. by saying the phase is blue-even, red-odd, we mean that we have reason to think it is; by saying the phase is blue-odd, red-even, we mean not that we have reason to think it is but merely that we have no reason to think the contrary.

But in general the definitions will have to be very complicated; we shall have, in order to verify that they are complete, to go through all the cases that satisfy the laws and consequences (together with any other propositions of the primary system we think right to assume) and see that in each case the definitions satisfy the axioms, so that in the end we shall come to something very like the general disjunctive definitions with which we started this discussion (p. 220). At best we shall have disjunctions with fewer terms and more coherence and unity in their construction; how much will depend on the particular case.

We could see straight off that (in a finite scheme) such definitions were always possible, and by means of \( \sigma \) and \( \tau \) we have reached no real simplification.

3. We have seen that we can always reproduce the structure of our theory by means of explicit definitions. Our next question is 'Is this necessary for the legitimate use of the theory?'
To this the answer seems clear that it cannot be necessary, or a theory would be no use at all. Rather than give all these definitions it would be simpler to leave the facts, laws and consequences in the language of the primary system. Also the arbitrariness of the definitions makes it impossible for them to be adequate to the theory as something in process of growth. For instance, our theory does not give any law for the colour of place 3; we should, therefore, in embodying our theory in explicit definition, define place 3 to be red unless it was observed to be blue (or else \textit{vice versa}). Further observation might now lead us to add to our theory a new axiom about the colour of place 3 giving say a cycle which it followed; this would appear simply as an addition to the axioms, the other axioms and the dictionary being unaltered.

But if our theory had been constructed by explicit definitions, this new axiom would not be true unless we changed the definitions, for it would depend on quite a different assignment of colours to place 3 at times when it was unobserved from our old one (which always made it red at such times), or indeed from any old one, except exactly that prescribed by our new axiom, which we should never have hit on to use in our definitions unless we knew the new axiom already. That is to say, if we proceed by explicit definition we cannot add to our theory without changing the definitions, and so the meaning of the whole.

[But though the use of explicit definitions cannot be necessary, it is, I think, instructive to consider (as we have done) how such definitions could be constructed, and upon what the possibility of giving them simply depends. Indeed I think this is essential to a complete understanding of the subject.]

4. Taking it then that explicit definitions are not necessary, how are we to explain the functioning of our theory without them?
Clearly in such a theory judgment is involved, and the judgments in question could be given by the laws and consequences, the theory being simply a language in which they are clothed, and which we can use without working out the laws and consequences.

The best way to write our theory seems to be this \( (\exists \alpha, \beta, \gamma) \) : dictionary . axioms.

The dictionary being in the form of equivalences.

Here it is evident that \( \alpha, \beta, \gamma \) are to be taken purely extensionally. Their extensions may be filled with intensions or not, but this is irrelevant to what can be deduced in the primary system.

Any additions to the theory, whether in the form of new axioms or particular assertions like \( \alpha(0, 3) \), are to be made within the scope of the original \( \alpha, \beta, \gamma \). They are not, therefore, strictly propositions by themselves just as the different sentences in a story beginning 'Once upon a time' have not complete meanings and so are not propositions by themselves.

This makes both a theoretical and a practical difference:

\( (a) \) When we ask for the meaning of e.g. \( \alpha(0, 3) \) it can only be given when we know to what stock of 'propositions' of the first and second systems \( \alpha(0, 3) \) is to be added. Then the meaning is the difference in the first system between \( (\exists \alpha, \beta, \gamma) : \text{stock} \cdot \alpha(0, 3) \), and \( (\exists \alpha, \beta, \gamma) \cdot \text{stock} \). (We include propositions of the primary system in our stock although these do not contain \( \alpha, \beta, \gamma \).)

This account makes \( \alpha(0, 3) \) mean something like what we called above \( \tau(\alpha(0, 3)) \), but it is really the difference between \( \tau(\alpha(0, 3) + \text{stock}) \) and \( \tau(\text{stock}) \).

\( (b) \) In practice, if we ask ourselves the question "Is \( \alpha(0, 3) \) true?", we have to adopt an attitude rather different from that which we should adopt to a genuine proposition.

For we do not add \( \alpha(0, 3) \) to our stock whenever we think
we could truthfully do so, i.e. whenever we suppose
(∃α, β, γ) : stock . α(0, 3) to be true. (∃α, β, γ) : stock .
α(0, 3) might also be true. We have to think what else we
might be going to add to our stock, or hoping to add, and
consider whether α(0, 3) would be certain to suit any further
additions better than α(0, 3). E.g. in our little theory either
β(n, 3) or β(n, 3) could always be added to any stock which
includes α(n, 3) . v . A(n) . B(n). But we do not add either,
because we hope from the observed instances to find a law
and then to fill in the unobserved ones according to that
law, not at random beforehand.

So far, however, as reasoning is concerned, that the values
of these functions are not complete propositions makes no
difference, provided we interpret all logical combination
as taking place within the scope of a single prefix (∃α, β, γ); e.g.

β(n, 3) . β(n, 3) must be (∃β) : β(n, 3) . β(n, 3),

not (∃β) β(n, 3) . (∃β) β(n, 3).

For we can reason about the characters in a story just as
well as if they were really identified, provided we don’t take
part of what we say as about one story, part about another.

We can say, therefore, that the incompleteness of the
‘propositions’ of the secondary system affects our disputes
but not our reasoning.

5. This mention of ‘disputes’ leads us to the important
question of the relations between theories. What do we mean
by speaking of equivalent or contradictory theories? or by
saying that one theory is contained in another, etc. ?

In a theory we must distinguish two elements:

(1) What it asserts : its meaning or content.

(2) Its symbolic form.

Two theories are called equivalent if they have the same
content, *contradictory* if they have contradictory contents, *compatible* if their contents are compatible, and theory $A$ is said to be *contained* in theory $B$ if $A$'s content is contained in $B$'s content.

If two theories are equivalent, there may be more or less resemblance between their symbolic forms. This kind of resemblance is difficult if not impossible to define precisely. It might be thought possible to define a definite degree of resemblance by the possibility of defining the functions of $B$ in terms of those of $A$, or conversely; but this is of no value without some restriction on the complexity of the definitions. If we allow definitions of any degree of complexity, then, at least in the finite case, this relation becomes simply equivalence. For each set of functions can be defined in terms of the primary system and therefore of those of the other secondary system *via* the dictionary.

Two theories may be compatible without being equivalent, i.e. a set of facts might be found which agreed with both, and another set too which agreed with one but not with the other. The adherents of two such theories could quite well dispute, although neither affirmed anything the other denied. For a dispute it is not necessary that one disputer should assert $\rho$, the other $\bar{\rho}$. It is enough that one should assert something which the other refrains from asserting. E.g. one says 'If it rains, Cambridge will win', the other 'Even if it rains, they will lose'. Now, taken as material implications (as we must on this view of science), these are not incompatible, since if it does not rain both are true. Yet each can show grounds for his own belief and absence of grounds for his rival's.

People sometimes ask whether a 'proposition' of the secondary system has any meaning. We can interpret this as the question whether a theory in which this proposition was denied would be equivalent to one in which it was affirmed.
This depends of course on what else the theory is supposed to contain; for instance, in our example $\beta(n, 3)$ is meaningless coupled with $\tilde{a}(n, 3) \vee \tilde{\gamma}(n)$. But not so coupled it is not meaningless, since it would then exclude my seeing red under certain circumstances, whereas $\tilde{\beta}(n, 3)$ would exclude my seeing blue under these circumstances. It is possible that these circumstances should arise, and therefore that the theories are not equivalent. In realistic language we say it could be observed, or rather might be observed (since 'could' implies a dependence on our will, which is frequently the case but irrelevant), but not that it will be observed.

Even coupled with $\tilde{a}(n, 3) \vee \tilde{\gamma}(n)$, $\beta(n, 3)$ might receive a meaning later if we added to our theory some law about the colour of 3. [Though then again $\beta(n, 3)$ would probably be a consequence of or a contradiction to the rest: we should then, I think, say it had meaning since e.g. $\beta(n, 3)$ would give a theory, $\tilde{\beta}(n, 3)$ a contradiction.]

It is highly relevant to this question of whether propositions have meaning, not merely what general axioms we include in our theory, but also what particular propositions. Has it meaning to say that the back of the moon has a surface of green cheese? If our theory allows as a possibility that we might go there or find out in any other way, then it has meaning. If not, not; i.e. our theory of the moon is very relevant, not merely our theory of things in general.

6. We could ask: in what sort of theories does every 'proposition' of the secondary system have meaning in this sense?

I cannot answer this properly, but only very vaguely and uncertainly, nor do I think it is very important. If the theory is to correspond to an actual state of knowledge it must contain the translations through the dictionary of many particular propositions of the primary system. These will, almost certainly, prevent many 'propositions' of the secondary
system from having any direct meaning. E.g. if it is stated in the theory that at time \( n \) I am at place 1, then for place 2 to be blue at that time \( n \) can have no direct meaning, nor for any very distant place at time \( n + 1 \). If then such 'propositions' are to have meaning at all, it must either be because they or their contradictories are included in the theory itself (they then mean 'nothing' or 'contradiction') or in virtue of causal axioms connecting them with other possible primary facts, where 'possible' means not declared in the theory to be false.

This causation is, of course, in the second system, and must be laid down in the theory.

Besides causal axioms in the strict sense governing succession in time, there may be others governing arrangement in space requiring, for instance, continuity and simplicity. But these can only be laid down if we are sure that they will not come into conflict with future experience combined with the causal axioms. In a field in which our theory ensures this we can add such axioms of continuity. To assign to nature the simplest course except when experience proves the contrary is a good maxim of theory making, but it cannot be put into the theory in the form 'Natura non facit saltum' except when we see her do so.

Take, for instance, the problem "Is there a planet of the size and shape of a tea-pot?" This question has meaning so long as we do not know that an experiment could not decide the matter. Once we know this it loses meaning, unless we restore it by new axioms, e.g. an axiom as to the orbits possible to planets.

But someone will say "Is it not a clear question with the onus probandi by definition on one side?" Clearly it means "Will experience reveal to us such a tea-pot?" I think not; for there are three cases:

(1) Experience will show there is such a tea-pot.
(2) Experience will show there is not such a tea-pot.
(3) Experience will not show anything.

And we can quite well distinguish (2) from (3) though the objector confounds them.

This tea-pot is not in principle different from a tea-pot in the kitchen cupboard.
B. GENERAL PROPOSITIONS AND CAUSALITY

Let us consider the meaning of general propositions in a clearly defined given world. (In particular in the common sense material world.) This includes the ordinary problem of causality.

As everyone except us¹ has always said these propositions are of two kinds. First conjunctions: e.g. ‘Everyone in Cambridge voted’; the variable here is, of course, not people in Cambridge, but a limited region of space varying according to the definiteness of the speaker’s idea of ‘Cambridge’, which is ‘this town’ or ‘the town in England called Cambridge’ or whatever it may be.

Old-fashioned logicians were right in saying that these are conjunctions, wrong in their analysis of what conjunctions they are. But right again in radically distinguishing them from the other kind which we may call variable hypotheticals: e.g. Arsenic is poisonous: All men are mortal.

Why are these not conjunctions?

Let us put it this way first: What have they in common with conjunctions, and in what do they differ from them? Roughly we can say that when we look at them subjectively they differ altogether, but when we look at them objectively, i.e. at the conditions of their truth and falsity, they appear to be the same.

\((x) . \phi x\) differs from a conjunction because

\(a\) It cannot be written out as one.

\(b\) Its constitution as a conjunction is never used; we never use it in class-thinking except in its application to a finite class, i.e. we use only the applicative rule.

¹ [I think that this must refer to himself and myself.—Ed.]
(c) [This is the same as (b) in another way.] It always goes beyond what we know or want; cf. Mill on 'All men are mortal' and 'The Duke of Wellington is mortal'. It expresses an inference we are at any time prepared to make, not a belief of the primary sort.

A belief of the primary sort is a map of neighbouring space by which we steer. It remains such a map however much we complicate it or fill in details. But if we professedly extend it to infinity, it is no longer a map; we cannot take it in or steer by it. Our journey is over before we need its remoter parts.

(d) The relevant degree of certainty is the certainty of the particular case, or of a finite set of particular cases; not of an infinite number which we never use, and of which we couldn't be certain at all.

(x) \( \phi x \) resembles a conjunction

(a) In that it contains all lesser, i.e. here all finite, conjunctions, and appears as a sort of infinite product.

(b) When we ask what would make it true, we inevitably answer that it is true if and only if every \( x \) has \( \phi \); i.e. when we regard it as a proposition capable of the two cases truth and falsity, we are forced to make it a conjunction, and to have a theory of conjunctions which we cannot express for lack of symbolic power.

[But what we can't say we can't say, and we can't whistle it either.]

If then it is not a conjunction, it is not a proposition at all; and then the question arises in what way can it be right or wrong.

Now in the case of a proposition right and wrong, i.e. true or false, occur doubly. They occur to the man who makes the proposition whenever he makes a truth-function of it, i.e. argues disjunctively about the cases of its truth and falsity.
Now this we never do with these variable hypotheticals except in mathematics in which it is now recognized as fallacious. We may seem to do so whenever we discuss the different theories obtainable by combining different proposed laws of nature. But here, if $P$ is such a law, we do not consider the alternatives $P$, i.e. $(x) . \phi x$, and $\bar{P}$, i.e. $(\exists x) . \phi x$, but we consider either having $P$ or not having $P$ (where not having it as a law in no way implies the law's falsity, i.e. $(\exists x) . \phi x$) or else having $P = (x) . \phi x$ or having $Q = (x) . \phi x$.

The other way in which right and wrong occur in connection with propositions is to an onlooker who says that the man's belief in the proposition is right or wrong. This, of course, turns simply on what the onlooker thinks himself and results from identity or difference between his view and what he takes to be that of the man he is criticising. If $A$ thinks $\rho$ and thinks also that $B$ thinks $\rho$, he says $B$ thinks truly; if he thinks $\rho$ and thinks that $B$ thinks $\bar{\rho}$, he says $B$ thinks falsely. But criticism may not always be of this simple type; it is also possible when $B$ thinks $\rho$, and $A$ thinks neither $\rho$ nor $\bar{\rho}$, but regards the question as unsettled. He may deem $B$ a fool for thinking $\rho$, without himself thinking $\bar{\rho}$. This happens almost always with hypotheticals. If $B$ says 'If I eat this mince pie I shall have a stomach-ache', and $A$ says 'No, you won't', he is not really contradicting $B$'s proposition—at least if this is taken as a material implication. Nor is he contradicting a supposed assertion of $B$'s that the evidence proves that so-and-so. $B$ may make no such assertion, in fact cannot always reasonably even if he is in the right. For he may be in the right without having proof on his side.

In fact agreement and disagreement is possible in regard to any aspect of a man's view and need not take the simple form of ' $\rho$ ', ' $\bar{\rho}$ '.

Many sentences express cognitive attitudes without being propositions; and the difference between saying yes or no
to them is not the difference between saying yes or no to a proposition. This is even true of the ordinary hypothetical [as can be seen from the above example, it asserts something for the case when its protasis is true: we apply the Law of Excluded Middle not to the whole thing but to the consequence only]; and much more of the variable hypothetical.

In order therefore to understand the variable hypothetical and its rightness or wrongness we must consider the different possible attitudes to it; if we know what these are and involve we can proceed easily to explain the meaning of saying that such an attitude is right or wrong, for this is simply having such an attitude oneself and thinking that one's neighbour has the same or a different one.

What then are the possible attitudes to the question—Are all men mortal?

(1) To believe it with more or less conviction.

(2) Not to have considered it.

(3) Not to believe it because it is unproven.

(4) Not to believe it because convinced that a certain type of man, who might exist, would be immortal.

(5) To disbelieve it as convinced that a particular man is immortal.

We have to analyse these attitudes; obviously in the first instance the analysis must be in terms of beliefs in singular propositions, and such an analysis will suffice for our present purpose.

To believe that all men are mortal—what is it? Partly to say so, partly to believe in regard to any $x$ that turns up that if he is a man he is mortal. The general belief consists in

(a) A general enunciation,

(b) A habit of singular belief.

These are, of course, connected, the habit resulting from