THE PSYCHOLOGICAL REVIEW

THE CONFLICTING PSYCHOLOGIES OF LEARNING —A WAY OUT¹

BY CLARK L. HULL Yale University

INTRODUCTION

One of the most striking things about the present state of the theory of learning and of psychological theory in general is the wide disagreement among individual psychologists. Perhaps the most impressive single manifestation of the extent of this disagreement is contained in 'Psychologies of 1925' (14) and 'Psychologies of 1930' (15). In these works we find earnestly defending themselves against a world of enemies, a hormic psychology, an act psychology, a functional psychology, a structural psychology, a Gestalt psychology, a reflexology psychology, a behavioristic psychology, a response psychology, a dynamic psychology, a factor psychology, a psychoanalytical psychology, and a psychology of dialectical materialism—at least a dozen.

No one need be unduly disturbed by the mere fact of conflict as such; that in itself contains an element of optimism, since it indicates an immense amount of interest and genuine activity which are entirely favorable for the advancement of any science. What disturbs many psychologists who are solicitous for the advancement of the science of psychology is

¹ The substance of this paper was read as a portion of the symposium on 'Psychological theories of learning,' at the Pittsburgh meeting of the A. A. A. S., December 28, 1934.

The writer is indebted to Dr. Robert T. Ross for the material appearing in notes 7 and 8. Dr. Ross has also read and criticized the entire manuscript. Professor Max Wertheimer and Dr. George Katona also read and criticized an early form of the manuscript.

that of which these disagreements are symptomatic. To put the matter in an extreme form: if all of these twelve psychologies should be in specific disagreement on a given point, then at least eleven of them must be wrong, and in such a welter of error the twelfth may very well be wrong also; at all events, it is difficult under such circumstances to see how all can be right about everything.

The obvious implication of this general situation has recently called out a timely little book by Grace Adams (1) entitled, 'Psychology: science or superstition?' In this work she points out what we all know only too well-that among psychologists there is not only a bewilderingly large diversity of opinion, but that we are divided into sects, too many of which show emotional and other signs of religious fervor. This emotionalism and this inability to progress materially toward agreement obviously do not square with the ideals of objectivity and certainty which we associate with scientific investigation; they are, on the other hand, more than a little characteristic of metaphysical and theological controversy. Such a situation leads to the suspicion that we have not yet cast off the unfortunate influences of our early associations with metaphysicians. Somehow we have permitted ourselves to fall into essentially unscientific practices. Surely all psychologists truly interested in the welfare of psychology as a science, whatever their theoretical bias may be, should cooperate actively to correct this.

But before we can mend a condition we must discover the basis of the difficulty. A clue to this is furnished by the reassuring fact that persisting disagreements among us do not concern to any considerable extent the results of experiment, but are confined almost entirely to matters of theory. It is the thesis of this paper that such a paradoxical disparity between scientific experiment and scientific theory not only ought not to exist but that it need not and actually will not exist if the theory is truly scientific. It will be convenient in approaching this problem first to secure a little perspective by recalling the essential characteristics of some typical scientific procedures.

FOUR TYPICAL SCIENTIFIC PROCEDURES

There are many approaches to the discovery of truth; for our present purposes these may be grouped roughly under four heads.

The simplest method of discovery is random observation —the trusting to chance that some valuable datum may turn up in the course of miscellaneous search and experiment. It is hardly conceivable that there ever will come a time in science when an experimenter will not need to be on the alert for the appearance of significant but unexpected phenomena. A classical example of the occasionally immense significance of such accidentally encountered observations is the discovery of the X-ray.

A second method of very wide and successful application in the search for truth is that sometimes known as systematic exploration. This seems to be the method advocated by Francis Bacon in his 'Novum Organum' (2). In modern times the discovery of salversan, by Ehrlich, illustrates in a general way this indispensable type of research procedure.

A third method widely employed in scientific investigations is that of the experimental testing of isolated hypotheses. Such isolated hypotheses often come as intuitions or hunches from we know not where; they occasionally appear in the form of prevailing traditions which are as yet inadequately tested by experiment. An example of the latter is the widespread belief that tobacco smoking interferes with the learning and thought processes (9).

A fourth procedure in the discovery of truth, and the one which particularly concerns us here, is found in experiments which are directed by systematic and integrated theory rather than by isolated and vagrant hypotheses. Such systematic theoretical developments are exemplified by relativity theory, chiefly in the hands of Einstein (7, 299), and by quantum theory (20), in the hands of a large number of individuals including Bohr, Rutherford, Heisenberg, Schrodinger, Dirac, and others. Perhaps the best-known investigation motivated by relativity theory is the astronomical observation whereby it was demonstrated that the image of a star whose light rays had passed close to the sun showed a certain amount of displacement from its true position, conforming both as to direction and amount with deductions made from the theory (7, 370). Possibly one of the most striking recent experiments based on quantum theory is the well-known discovery and isolation of 'heavy' water, at Columbia University a few months ago, by Professor Urey.

Our special concern here is to point out that this fourth type of investigation, in addition to yielding facts of intrinsic importance, has the great virtue of indicating the truth or falsity of the theoretical system from which the phenomena were originally deduced. If the theories of a science really agree with the experimental evidence, and if there is general agreement as to this evidence, there *should* be a corresponding agreement regarding theory. An examination of the nature of scientific theoretical systems and their relationship to the fourth type of scientific procedure just considered should aid us in coping with the paradox presented by the present unfortunate state of psychological theory.²

FOUR ESSENTIALS OF SOUND SCIENTIFIC THEORY

It is agreed on all hands that Isaac Newton's 'Principia' is a classic among systematic theories in science. It starts with eight explicitly stated definitions and three postulates (laws of motion) (16, pp. 1–13), and from these deduces by a rigorous process of reasoning the complex structure of the system. Many persons who may not be overly familiar with the technical details of classical mathematical physics will be able to understand the essentials of such a system from our knowledge of ordinary Euclidian geometry, which as a systematic structure is substantially similar. In the geometries we have our definitions, our postulates (axioms), and, following these, the remarkable sequence of interrelated and inter-

² This emphasis on the fourth type of experimental approach is not to be understood as an advocacy of it as an exclusive method in psychology; neither is it being urged that theoretical considerations are paramount. Many approaches are necessary to produce a well-rounded science. Some temperaments will prefer one approach, some another, thus leading to a useful division of labor. locking theorems which flow so beautifully by deduction from the basic assumptions. In a truly scientific system, however, a considerable number of the theorems must constitute specific hypotheses capable of concrete confirmation or refutation. This was eminently true of Newton's system. For a very long time the Newtonian physics stood this test, though finally certain important deductions from his postulates failed of confirmation, and it fell. Had Newton's system not been firmly anchored to observable fact, its overthrow would not have been possible and we would presumably be having at the present time emotionally warring camps of Newtonians and Einsteinians. Fortunately, we are spared this spectacle.

To summarize in a formal and systematic manner, it may be said that for a candidate to be considered as a sound scientific theory it must satisfy four basic criteria.³

I. The definitions and postulates of a scientific system should be stated in a clear and unambiguous manner, they should be consistent with one another, and they should be of such a nature that they permit rigorous deductions.

II. The labor of deducing the potential implications of the postulates of the system should be performed with meticulous care and exhibited, preferably step by step and in full detail. It is these deductions which constitute the substance of a system.

III. The significant theorems of a truly scientific system must take the form of specific statements of the outcome of concrete experiments or observations. The experiments in question may be those which have already been performed, but of particular significance are those which have not previously been carried out or even planned. It is among these latter, especially, that crucial tests of a theoretical system will be found.⁴

⁸ As the reader examines these items it might be illuminating for him to consider the particular theoretical system which is his special aversion, and judge whether or not it passes each successive criterion. After having thus fortified himself, he might proceed cautiously to a similar examination of the system which he favors.

⁴ For this reason it is especially desirable for the advancement of science that the proponents of theoretical systems publish the deductions of the outcome of as yet untried experiments. The failure of subsequent experimental verification of such deIV. The theorems so deduced which concern phenomena not already known must be submitted to carefully controlled experiments. The outcome of these critical experiments, as well as of all previous ones, must agree with the corresponding theorems making up the system.

Let us consider briefly some of the more important reasons why a sound scientific system should possess these four characteristics. Consider the first: If the postulates of an alleged system are not stated clearly they can hardly be known to the scientific public which may wish to evaluate the system. Moreover, if the postulates have never been explicitly written out by the sponsor of the system, the chances are high that they are not clear even to him. And, obviously, if the definitions and postulates of a system are not clear to the sponsor of the system, neither he nor anyone else can make specific and definite deductions from them.

Second, deductions must be performed with rigor because only in this way can their implications become known. Obviously, until the implications of the postulates are known they cannot possibly be submitted to experimental test; and unless the deductions are rigorous the experimental test will be futile because it will have no real bearing on the soundness of the postulates. Indeed, without rigorous deductions a would-be system is nothing more than a vague and nebulous point of view.

Third, the deductions must be related specifically to the concrete data of the science in question, since otherwise they cannot be submitted to the absolutely indispensable experimental test. It is here that scientific theory differs (or *should* differ) sharply from metaphysical speculations such as concern ethics and theology. Metaphysics does not permit this continuous check on the validity of the deductions, which largely accounts for the interminable wrangles characteristic of that literature. This criterion accordingly becomes in-

ductions should not be regarded as in any way discrediting the author. Instead, it should be considered merely a normal incident in the evolution of science. Fortunately, in such situations it is the hypothesis which is on trial, not the proponent's reputation as a seer.

valuable in distinguishing psychological metaphysics from scientific psychological theory. By this criterion much of what at present passes as theory in our literature must be regarded as metaphysical, *i.e.*, as essentially unscientific.

Fourth, the labor of setting up the critical experiments designed to verify or refute the theorems thus rigorously deduced from the postulates must be performed thoroughly and impartially because, once more, we shall otherwise lack the indispensable objective test of the truth of the system.

It scarcely needs to be added that there is nothing either radical or new in the above criteria of sound scientific theory; on the contrary, the program is conservative and respectable to an eminent degree. Indeed, it has been accepted in science for at least two hundred years. Our purpose is mainly to urge that we really put into practice what we, with the other sciences, have known for a very long time. This we evidently have not done; otherwise we would not be confronted with the glaring paradox of the wildest confusion in the matter of theory coupled with substantial agreement in the field of experiment.

IS RIGOROUS THEORY IN PSYCHOLOGY POSSIBLE?

No doubt many will feel that such standards of scientific theory may be suitable for theoretical physics, but that they are quite impossible in psychology, at least for the present. To take such a view is equivalent to holding that we can have no genuinely scientific theory in psychology. This is indeed conceivable, but if so we ought not to pretend to have theories at all. If scientific theories are really impossible in psvchology, the quicker we recognize it, the better. There are signs, however, that the beginnings of a genuinely scientific theory of mammalian behavior are already on their way. Extremely promising examples of such achievements in intimately related fields have been published by Crozier (3) and by Hecht (8). The recent work of Gulliksen (6), in which he presents a genuinely rational equation for the learning curve, as distinguished from an empirically fitted formula, offers promise of a larger development in the field of mammalian learning.

It is probably not accidental that all three of the above studies are essentially mathematical. At present, on the other hand, the superficial appearance of the concepts regarding learning which are current among our theorists does not suggest ready mathematical treatment. And while this condition is probably more apparent than real, it serves to raise the important question as to whether rigorous logical deductions can be made on the basis of such quasi-mathematical concepts as have so far emerged from behavior experiments.

There is reason to believe that a genuinely scientific system may be constructed from such materials, and that the difficulty of making such theoretical constructs is not nearly so great as their rarity might lead one to expect. Obviously, the best evidence for such a belief is actual performance. Accordingly, the following section (pp. 501 ff.) of this paper is given over to the presentation of a suggested miniature scientific system based on typical quasi-mathematical concepts. This has been developed by means of a form of reasoning analogous to that employed in ordinary geometrical proofs. In it an effort has been made to conform to the criteria laid down above as necessary for a sound theoretical development. It is hoped that it will aid in making clear in some concrete detail the theoretical methodology here being advocated. Let us, accordingly, proceed to the critical examination of this miniature theoretical system in the light of our four formal criteria of what scientific theory should be.

At the beginning (pp. 501 ff.) there will be found a series of eleven definitions: of rote series, of the learning of rote series, of excitatory tendency, of inhibitory tendency, of spanning, of actual and of effective strength of excitatory tendencies, of remote excitatory tendency, of trace conditioned reaction, and so on.

Next there appears (p. 503) a series of explicitly stated postulates: that the remote excitatory tendencies of Ebbinghaus exist; that remote excitatory tendencies of Ebbinghaus possess the same behavior characteristics as do the trace conditioned reflexes of Pavlov (Lepley's hypothesis); that the period of delay of trace conditioned reflexes possesses an inhibition of delay; that inhibitions are additive; that caffeine retards the accumulation of inhibition; that inhibitions diminish more rapidly with the lapse of time than do related excitatory tendencies, and so on. So much for the first criterion.

There follows (pp. 504 ff.) a series of eleven theorems derived by a formal process of reasoning from the preceding postulates and definitions. For the most part each step of the reasoning is explicitly stated and the logical source of each is conscientiously given. In this connection it is to be observed that the deduction or proof of each theorem is a complex multiple-link logical construct involving the joint action of numerous principles or postulates, as contrasted with simple syllogistic reasoning where but two premises are employed. Moreover, it is to be noted that the system is an integrated one not only in that all the theorems are derived from the same postulates, but also in that the later theorems are dependent on the earlier ones in the form of a logical hierarchy, very much as in systems of geometry. In the derivation of these eleven theorems an attempt has thus been made to conform to the second criterion of a satisfactory scientific system.

Let us now proceed to the examination of this theorem hierarchy from the point of view of the third and fourth criteria.

The first four theorems, while logically necessary for the derivation of the later ones, do not themselves permit any direct experimental test. It is believed, however, that all of the others are sufficiently concrete and specific to permit definite experimental confirmation or refutation. Consider, for example, Theorem V. In plain language, this states that the central portion of a rote series is more difficult to memorize than are the two ends. This is, of course, a fact long known to experimentalists (21).⁵ Theorem VI, which states that

⁵ It is to be noted, however, that while the general picture of series difficulty as shown by experiment agrees with the theorem, there is disagreement in detail. The theorem demands that the maximum difficulty appear in the exact center of the series, whereas it actually appears a little posterior to the center. This, of course, reflects an inadequacy in the theory and calls for a revision of postulates. This systematic reconstruction has already gone far enough to correct the difficulty here considered. This the difficulty of learning syllables increases most rapidly at the ends of the series but the rate of increase is less and less as the point of maximum difficulty is approached, has also long been a laboratory commonplace (21). Theorem VII states that the reaction times of the syllables of a rote series will be shortest at the ends and progressively longer as the middle is approached; this is a case of a deduction actually made in advance of experiment. Recently, however, the deduction has had experimental confirmation (24).

Now, let us look at Theorem VIII. This theorem means that syllables in the middles of partially learned series are known better a short time after the termination of practice than they are immediately at the conclusion of practice. It is particularly to be noted that this theorem flies directly in the face of the old and time-honored principle of forgetting; *i.e.*, it demands that performance shall *improve* instead of deteriorate with the passage of time. When this deduction was first performed our logic seemed to be carrying us into a topsy-turvy world, but our postulates presented us with no alternative; scientific theory is concerned with inflexible logic rather than with predictions based on intuitions or wishes. A year or two after the deduction was made, Ward submitted it to critical experimental test and found the theoretical expectation fully and completely substantiated (24).

And so we could go on through Theorems IX and X. It will suffice to say that Theorem IX has recently been experimentally verified by Ward (24) after the deduction was made, and that Theorem X states a striking law of economy of learning long known to the literature (18, 375 f.).

Finally we come to Theorem XI. Stripped of technical verbiage, this theorem means that the peak of difficulty in the middle of a rote series when learned by massed practice under the influence of caffeine will be lower than when learned by massed practice in the normal condition. Two or three years after this deduction had been made, the author set up an experiment may serve as an example of the successive-approximation procedure characteristic of theoretical development in science. The revised construct will be given in connection with a full statement of the system to be contained in a contemplated publication. especially to test it. When the experiment was completed and the data tabulated, it was found that the deduction was not verified—the peak of difficulty in the middle of the series was a little higher under caffeine than in the control series, where the subjects learned the material in a normal condition (10). Here, then, is a case where a definite deduction has been flatly controverted by fact.

Clearly, where a theory is opposed by a fact, the fact has the right of way. In a situation of this kind something is obviously wrong, presumably with one or more of the postulates involved in the deduction. In this particular case suspicion naturally rests most heavily on Postulate VI. At all events, Theorem XI serves to round out and give a further note of realism to this miniature scientific theoretical system. It is a noteworthy event, in the present status of psychological theory, to have a deduction sufficiently anchored by logic to the postulates of the system that a collision with a stubborn experimental fact shall be able to force a revision of the system. It is reasonably safe to assume that the rarity of such collisions at present is not due to the infallibility of current theoretical constructs. Until our systems become sufficiently clear and definite for this kind of event to be of fairly frequent occurrence, we may well suspect that what passes as theory among us is not really making contact with our experimental facts.

A MINIATURE SCIENTIFIC THEORETICAL SYSTEM BY WAY OF ILLUSTRATION

Definitions

I. A rote series is a number of nonsense syllables presented visually one at a time for constant periods (e.g., three seconds) with only a fraction of a second between exposures. The subject learns to speak each syllable while its predecessor is still in view, the overt immediate stimulus for each overt reaction being the visual stimulus arising from the preceding syllable.

II. A rote series is said to be learned when the subject can correctly anticipate each successive syllable throughout a single repetition.

III. An 'excitatory tendency,' as emanating from a stimulus, is a tendency for a reaction to take place more certainly and, in case it does occur, to do so more vigorously other things equal, soon after the organism has received said stimulus than at other times.

IV. An 'inhibitory tendency' is one which has the capacity to weaken the action potentiality of a concurrent excitatory tendency.

V. A syllable reaction tendency is said to be spanned by a remote excitatory

tendency and by the parallel inhibition of delay (Postulate III) when said syllable reaction tendency falls between the stimulus syllable and the response syllable associated with the remote excitatory tendency and the parallel inhibition of delay in question.

VI. The 'actual' strength of an excitatory tendency is that strength it would display for action if uncomplicated by concurrent inhibitory tendencies.

VII. The 'effective' strength of an excitatory tendency is that strength it displays in action under whatever conditions of inhibition may exist at the time.

VIII. A remote excitatory tendency is an excitatory influence, initiated by a syllable as a stimulus, exerted upon any other syllable as a reaction with the exception of the syllable immediately following the stimulus syllable.

IX. A trace conditioned reaction is an $S \rightarrow R$ relationship (acquired in isolation by a special conditioning technique) which has the characteristic that an appreciable interval (*e.g.*, sixteen seconds) may elapse between the presentation of the overt stimulus and the taking place of the overt response.⁶



FIG. 1.

Diagrammatic representation of both the immediate and the remote forward excitatory tendencies assumed to be operative in rote series. The straight broken arrows represent immediate excitatory tendencies and the curved solid arrows represent remote excitatory tendencies. The number of remote excitatory tendencies spanning a given syllable, such as ZIT, is given by the formula (n - 1) (N - n) where N is the total number of syllables in the series and n is the ordinal number of the syllable whose span value is under consideration. Thus, in the above example, N = 7 and the n for ZIT = 3. Accordingly, n - 1 = 2 and N - n = 4. Consequently, ZIT should have 2×4 or 8 remote excitatory tendencies spanning it. The truth of this computation may be verified by counting the number of curved lines immediately above the syllable in question. The number of remote excitatory tendencies spanning the several syllables is given beneath each.

⁶What is spoken of as the 'overt' stimulus of a trace conditioned reaction is not regarded as the 'actual' stimulus. The 'overt' stimulus is supposed to set in motion some kind of slowly changing internal sequence more or less characteristic of each such stimulus. It is the stimulus value of the phase of this sequence immediately preceding the reinforcing stimulus which is regarded as the 'actual' stimulus of the trace conditioned reaction. It thus comes about that the stimulus of POF (Fig. 1) is compounded of 6 elements from as many different sources, whereas that of FAP arises from a single source. But, so far as is now known, the ease of conditioning is not influenced by the complexity of the stimulus, so that the 'actual' strength of the excitatory tendencies to the arousal of POF and FAP should be alike so far as this factor is concerned. This means, necessarily, that the *immediate* excitatory tendency from KEX to POF must be appreciably weaker than that from KEM to FAP or even from ZIT to YEV. This last deduction is obviously capable of experimental test. X. 'Massed practice' is a method of learning in which the series is run through from beginning to end almost continuously, *i.e.*, with a pause only of from ten to twenty seconds between successive repetitions.

XI. 'Distributed practice' is a method of learning in which an appreciable interval of time (e.g., one hour or more) is interposed between successive repetitions; otherwise it is the same as 'massed practice.'

Postulates

I. Rote series possess functionally potent remote excitatory tendencies extending forward from each syllable of the series as an overt stimulus to every syllable placed later in the series as an overt response except the response syllable immediately following the stimulus syllable. (Ebbinghaus, 4, 106.)

II. The remote excitatory tendencies of Ebbinghaus possess the same characterstics as the trace conditioned reflexes of Pavlov. (Lepley's hypothesis, 12; 13.)

III. The period of delay of trace conditioned reflexes possesses a power to inhibit (temporarily) to a certain extent the functional strength of excitatory tendencies, the reactions of which would otherwise tend to take place during such period. (Pavlov, 17, 173.)

IV. The inhibition of delay of each succeeding degree of remoteness (distance between overt stimulus and overt response) decreases progressively, each additional increment in remoteness diminishing the inhibition, on the average, by a constant amount. (Assumed by rough analogy to corresponding excitatory tendencies, 4, 106.)

V. Inhibitions of delay operative at the same time summate arithmetically. (Assumed from analogy to excitatory tendencies, 22, 36 f.)

VI. Inhibitions of delay accumulate to a lesser degree when the subject is under the influence of caffeine than do associated excitatory tendencies. (Evans, 5, 365.)

VII. When learning is performed by massed practice, the ratio of the actual strength of excitatory tendency to the inhibition of delay is, on the average, constant throughout the learning process, and such as usually to leave a positive effective strength of excitatory tendency. (Assumed as a first approximation.)

VIII. Inhibitory tendencies in the early stages of weakening through the lapse of time diminish more rapidly than do associated excitatory tendencies. (Pavlov, 17, 99 and 58 f.)

IX. A constant minimal strength of excitatory tendency is necessary to make recall possible even when no concurrent inhibition is present. (Assumed.)

X. The total aggregate actual excitatory tendency exerted on a syllable as a reaction tendency is, on the average, a constant for all syllables in a given list at a given time. (Assumed.)

XI. A constant minimal 'effective' strength is required of any given excitatory tendency for it to pass the threshold of overt reaction. (Assumed.)

XII. Under the conditions of rote learning, each repetition of a rote series adds, on the average, a constant positive increment to the actual strength of each excitatory tendency of the series. (Pillsbury, 18, 370.)

XIII. The greater the functional or 'effective' strength of the excitatory tendency evoking a reaction, the less, on the average, will be the time elapsing between the stimulus and the reaction. (Simley, 23.)

XIV. The 'actual' strength of excitatory tendencies accumulated through repetitions is not influenced by the previous presence of superposed inhibitions of delay. (Assumed.)

Theorems

I

If the number of syllables in a rote series is N, and the ordinal number of a particular syllable counting from the beginning is n, the syllable as a reaction tendency will be spanned by (n - 1) (N - n) remote excitatory tendencies.

I. It is evident (Postulate I and Fig. 1) that a given syllable in a rote series (Definition I) is spanned (Definition V) by remote excitatory tendencies (Definition VIII) all of which originate in the syllables anterior to itself and which terminate in syllables posterior to itself; *i.e.*, each syllable anterior to a given syllable has a remote excitatory tendency extending to each syllable posterior to said syllable *n*.

2. Since there are (n - 1) syllables anterior to a given syllable and (N - n) syllables posterior to it, it follows from (1) and Postulate I that there must be (n - 1) (N - n) remote excitatory tendencies spanning any given syllable as a reaction.

II

Within any rote series, the mean degree of remoteness of remote excitatory tendencies spanning a given syllable is the same for all syllables, viz., $\frac{N+1}{2}$.

1. In continuous series the terms of which increase by constant steps, the mean of the series as a whole will be given by the mean of the values appearing at the respective ends of the series.

2. By Postulate I (and Fig. 1), the remote excitatory tendencies spanning a given syllable and originating in a particular syllable, satisfy the conditions of (1).

3. Take any syllable, n, of a rote series. It is evident (Fig. I and Postulate I) that those remote excitatory tendencies originating in syllable I and which span syllable n must have as their greatest length the distance in intervals from the last syllable of the series to the first syllable of the series, *i.e.*, N - I intervals, and for their shortest value the distance in intervals from syllable I to syllable n + I, *i.e.*, n + I - I, or simply n intervals.

4. From (1), (2), and (3) it follows that the remote excitatory tendencies of the set emanating from syllable 1 have as their mean that of N - 1 and n, or $\frac{N + n - 1}{2}$.

5. It is evident also (Postulate 1 and Fig. 1), that the excitatory tendencies of the set emanating from the second syllable must all be one step less in distance than those emanating from syllable 1, *i.e.*, that their mean value must be $\frac{N+n-1}{2} - 1$; that the mean of those emanating from syllable 3 must be $\frac{N+n-1}{2} - 2$, and so on, the amount subtracted from the fraction in the case of the mean of the last set being one less than the total number of sets.

6. But by (2) of Theorem I, the number of such sets is n - 1. It follows from (5) that the value subtracted from the fraction which appears in the formula representing the mean of the last set must be (n - 1) - 1, or n - 2.

7. From (4), (5), and (6), the final mean of the series must be $\frac{N+n-1}{2}$ - (n-2). But by (5) and (6) the means of the several series constitute a continuous series exhibiting constant step intervals. Therefore, by (1), the mean of these means must be given by the mean of the first and last means of the series. 8. By (5), (6), and (7), the mean extent of the series of means must be

$$\frac{N+n-1}{2} + \frac{N+n-1}{2} - (n-2)$$

which becomes

$$\frac{N+n-1+N+n-1-2n+4}{4}$$

The n's disappear, leaving

$$\frac{2N+2}{4} \quad \text{or} \quad \frac{N+1}{2}.$$

9. But since by assumption n was any syllable, it follows from (7) that the mean length of remote excitatory tendencies spanning any syllable is like that of all the others, viz., $\frac{N+1}{2}$.

ш

The total inhibition of delay operative at any given syllable position is measured by the number of remote excitatory tendencies spanning that syllable position.

1. By Postulates II and III and Definition IX, the intervals of delay of remote excitatory tendencies are the loci of inhibitions of delay.

2. By Postulate IV, the magnitude of these inhibitions of delay is a decreasing linear function of the degree of remoteness of the excitatory tendency in question.

3. It follows from (1) and (2) and Theorem II that the mean magnitude of inhibition (Definition IV) effective at any given syllable position in the series must be like that of all other syllable positions.

4. But if the mean inhibition of delay at all syllable positions is the same, it follows that the total inhibition at any given syllable position must be strictly proportional to the number of remote excitatory tendencies spanning that syllable position.

5. From (4) and Postulate V the theorem follows.

IV

The number of repetitions required for mastery of any particular syllable of a rote series is $T + R_I$, where T is a constant representing the number of repetitions required to produce learning when no inhibition is present, and R_I is a linear function of the number of spannings, i.e. of (n - 1) (N - n).

1. By Postulates IX and XII and Definition II, a finite basic number of repetitions, T, will be required to produce the strength of excitatory tendency (Definition III) necessary to evoke reaction when there is no inhibition present.

2. By Postulate X and Definition VI, T must be a constant throughout any given rote series.

3. By Postulates XI and XII and Definitions IV and VII, there must be added to the threshold constant, T, certain repetitions to overcome any inhibitions present.

4. By Postulates V and XII, the number of repetitions at a given syllable will be a direct linear function of the aggregate inhibition at that syllable.

5. By (4) and Theorem III, the number of repetitions required to override the inhibition at any point within a given series must be a linear function of the value (n - 1) (N - n).

6. From (2) and (5) it follows that the number of repetitions required for mastery of a rote series at any given point must be the sum of those required to pass the thres-

hold of recall, T, plus those required to overcome the adverse influence of inhibition, (n - 1) (N - n), *i.e.*, it must be $T + R_I$ where the latter is a linear function of (n - 1) (N - n). V

The number of repetitions required for mastery of the individual syllables of a rote series is greater in the central region of the series than at either end, the position of maximum difficulty falling at point $\frac{N+1}{2}$.

I. Since, by Theorem IV, T in the expression $T + R_I$ is a constant, it follows that the variability in the number of repetitions required for the mastery of the several portions of a rote series will be a direct linear function of (n - 1) (N - n) only, since R_I is a linear function of (n - 1) (N - n).

2. If, now, we substitute in this formula the successive ordinal values at the beginning of any rote series, taking the length of the series at any convenient value such as N = 9, we have,

Syllable number (n),	I	2	3	4	5	6	7	8	9
Units of repetition to learn,	0	7	12	15	16	15	12	7	c

3. It may be seen by an inspection of the series of values in (2) that the number of repetitions required for mastery increases continuously from the ends toward the middle of the series, the maximum falling at point 5, which may be expressed by $\frac{N+1}{2}$. Thus we have a concrete demonstration of the truth of the theorem for a

particular series.7

VI

The rate of increase in the number of repetitions required for mastery in a rote series progressively diminishes as the point of maximal difficulty is approached from either end, I. Taking any convenient length of series such as one of eight syllables (N = 8),

⁷A deduction of the essential portion of this theorem is yielded by the calculus:

$$R_I = a + m(n-1)(N-n)$$

Expanding we have,

$$R_{I} = a - m[+ n^{2} + n(N + 1) - N]$$

Differentiating with respect to n,

$$\frac{dR_{\rm I}}{dn} = m[-2n + (N+{\rm I})]$$

at the maximum,

$$\frac{dR_I}{dn} = 0$$

whence

-2n+(N+1)=0

and solving for n we have,

$$n=\frac{N+1}{2}$$

therefore, the position of maximum difficulty falls at the point $\frac{N+1}{2}$.

we have by Theorem IV the formula $T + R_I$, remembering that T is constant and R_I is a linear function of (n - 1)(N - n).

Syllable number (n),	I	2	3	4	5	6	7	8
Units of repetition to learn,	0	6	10	12	12	10	6	о

2. Here it may be seen that the units of repetition required for mastery increase by 6 points from syllable 1 to syllable 2, by 4 points from syllable 2 to syllable 3, and by 2 points from syllable 3 to syllable 4; *i.e.*, the rate of increase in difficulty progressively diminishes as the middle is approached.

3. A corresponding inspection reveals the same type of progression from the postterior end of the series as point $\frac{N+1}{2}$ is approached.

4. (1), (2), and (3) constitute a concrete demonstration of the truth of the theorem for a particular series.⁸

VII

The reaction times of the syllables of a rote series learned by massed practice will be shortest at the end positions and progressively longer the farther the syllable from the ends of the series.

1. By Theorem V, syllables require an increasing number of repetitions to learn as the point of maximal difficulty of the series is approached from either end.

2. From (1) and Postulates XI and XII and Definition VII, it follows that the syllables near the ends of the series will rise above the threshold of recall progressively earlier than the syllables farther from the ends.

⁸ A deduction of the essential portion of this theorem is yielded by the calculus (see note to Theorem V): It follows from

$$\frac{aKI}{dn} = m[-2n + (N+1)] \text{ (where } m \text{ is positive)}$$

that

$$\frac{d^2R_I}{dn^2} = -2m$$

whence, if

$$n < \frac{N+1}{2}$$

 $\frac{dR_I}{dn} \quad \text{is positive} \quad \text{whence, the curve increases toward the right with decreasing} \\ \frac{d^3R_I}{dr^2} \quad \text{is negative} \quad \text{slope} \quad \text{slope}$

if

$$n > \frac{N+1}{2}$$

 $\frac{dR_I}{dn} \quad \text{is negative} \\ \frac{d^2R_I}{dn^2} \quad \text{si negative} \end{bmatrix} \text{ whence, the curve decreases toward the right with increasing}$ (negative) slope.

3. From (2), Definition I, and Postulate XII, it follows that the syllables near the ends of the series will be overlearned more than those at the middle, *i.e.*, they will have progressively stronger effective excitatory tendencies (Definition VII) as their distance from the middle of the series increases.

4. By (3) and Postulate XIII the theorem follows.

VIII

In rote series learned to a variable but incomplete degree by massed practice, the number of successful reactions in the middle portion of the series will be greater after a certain period of no practice than at once after the conclusion of learning.

1. By Theorems I and III and Postulate VII, it follows that throughout the learning of rote series where the learning is performed by massed practice there will be variable but finite amounts of inhibition operative on the excitatory tendencies of syllables in the interior of series, *i.e.*, upon all but the two end syllables.

2. By Definition IV, this will depress the effective reactive capacity of such excitatory tendencies (Definition VII) below their actual values.

3. But, by Postulate VIII, inhibitions at first diminish more rapidly during the passage of time than do the associated excitatory tendencies.

4. By (3), during a given interval of no practice the inhibitory tendency will decrease by a finite amount.

5. It follows from (2) and (3) and Postulate XIV that in the early stages of a period of no practice following the learning of a rote series, the effective excitatory strengths of the interior syllables as reaction tendencies will be greater by finite amounts than at the conclusion of learning.

6. From (5) it follows that all syllables as reaction tendencies whose excitatory strengths are above the reaction threshold at the conclusion of incomplete learning will remain above after the period of no practice.

7. Since the degree of learning before interruption varies from one series to another (as here assumed), it follows that of those reaction tendencies which are below the threshold of recall some will differ from the threshold by an amount less than the finite amount indicated in (4).

8. From (3) and (7) it follows that certain syllables which are below the threshold of recall at the conclusion of incomplete learning will be above it at the conclusion of an optimal interval early in the period of no practice.

9. The group of effective reaction tendencies above the threshold at the conclusion of learning (6) added to the group which pass the threshold after an optimal interval of no practice (8) will make a sum larger than the former alone, from which the theorem follows.

IX

In just barely learned rote series the reaction time of syllables in the interior of the series will be shorter after an optimal period of no practice than for the corresponding individual syllables at the conclusion of learning by massed practice.

1. By reasoning analogous to that of (1), (2), (3), and (4) of the proof for Theorem VIII, it follows that the effective excitatory strength of just-learned syllables in the middle of rote series will be greater at some point early in the period of no practice than at the conclusion of learning by massed practice (Definition X).

2. By (1) and Postulate XIII, this increased excitatory strength will be accompanied by shortened reaction time, from which the theorem follows. Rote series will be learned with fewer repetitions by distributed practice than by massed practice.

1. By Theorems II, III, and IV, the most difficult syllables to memorize of a rote series are loaded with inhibitions of delay.

2. By Definition XI, the method of distributed practice involves appreciable periods of time between repetitions. By Postulate VIII these time intervals, if not too long, will dissipate the inhibition more rapidly than the associated excitatory tendency. It follows that for a given amount of training the method of distributed practice will yield relatively less accumulated inhibition than by massed practice.

3. From (2) it follows (Postulates XI and XII) that the method of distributed repetitions will bring the most difficult syllable above the threshold of recall with fewer repetitions than will be the case by the method of massed repetitions.

4. But, by Definitions I and III, the number of repetitions required to learn rote series is that required to learn the most difficult single syllable.

5. By (3) and (4), the theorem follows.

XI

The value obtained by dividing the number of repetitions required to bring syllables above the threshold at the ends of rote series, by the number required in the middle of the same series, will be larger when the learning is done under the influence of caffeine than when done in the normal condition, the learning in both cases to be performed by massed practice.

1. By Theorem V, the middles of rote series learned by massed practice require more repetitions for learning than do the ends.

2. From (1) it follows that the number of repetitions per syllable for learning at the ends divided by the number at the middle R_E/R_M will yield a value less than 1.

3. Now, by Postulate VI, inhibitions accumulate to a lesser degree, other things equal, when the learning is performed under the influence of caffeine. It follows from this and Theorems II and III that less inhibition will accumulate in the middle of the series in question when learning is performed under the influence of caffeine.

4. By (3), Definition IV, and Postulate XII, it follows that the middle syllables will be learned with less repetitions under caffeine than in the normal condition, *i.e.*, that R_M will be smaller than normal. Since caffeine has no such influence on syllables not inhibited, R_K will remain the same.

5. But to reduce R_M in the division R_E/R_M will increase the resulting values.

6. From (5) the theorem follows.

Some Problems Connected with the Evaluation of Psychological Theory

The recognized principles of science, then, provide us with a method which seems capable of bringing some kind of order out of the present chaos in theoretical psychology. Moreover, the program appears to be one to which all theorists, however diverse their postulates provided they are not essentially metaphysical or mystical, may subscribe. Indeed, it seems to be so firmly rooted in the traditions and essential logic

509

of science that all would-be theoretical work will ultimately come to be judged by the scientific public according to this standard, regardless of the views of the theorists themselves. This brings us to the consideration of certain concrete problems which arise when an attempt is made to evaluate the claims of competing theoretical systems.

In the first place, it should be obvious that all mere systems of classification must be rejected. A dictionary may be systematic, but it can hardly be rated as a theoretical system even when the terms are largely of new coinage. Merely to call a bit of learning behavior a case of 'closure' or 'insight' on the one hand, or a case of 'conditioning' or 'trial-anderror' on the other, will not serve. Such systems cannot pass even the first criterion.⁹

Next we must consider the nature of the concepts and postulates which are admissible as the basis for psychological theory. Some psychologists appear to have assumed that only principles incapable of direct observational verification ¹⁰ should be admitted as postulates, whereas others may conceivably have assumed that only principles capable of direct observational verification should be admitted. In a similar manner, one group of theorists may insist that the postulates from which psychological systems evolve must be concerned with *parts*, while another group may insist that they must concern *wholes*. One group of theorists may insist that they postulates must come solely from conditioned reflex experiments, whereas to another group such postulates might not be at all acceptable.

⁹ It appears to be at this point that most current attempts in the field of psychological theory break down. Their concepts appear not to be of such a nature that significant theorems may be drawn from them by a rigorous logic. A theoretical system without proven theorems is a paradox, to say the least.

¹⁰ The postulates of a system may be susceptible of two types of verification—one indirect and the other direct. Indirect verification occurs when a deduction from a combination of postulates is observationally confirmed. The failure of such a verification throws doubt on the soundness of all of the postulates involved. This particular doubt is removed when appropriate change is made in one or more of these postulates so that deductions from them conform not only to the new observations but to all those phenomena previously deduced and verified. All postulates are susceptible of indirect verification, but some postulates permit direct verification and some do not. Postulates regarding the positions and movements of electrons, for example, permit indirect verification but not direct observation. From the present point of view this argument is quite footless. Actually, all such groups beg the main question. The question at issue is: Can more theorems which will be confirmed in the laboratory be deduced from postulates which are principles of dynamics, or more from postulates which are principles of mechanics, or more from a combination of both types of postulates; can more sound theorems be deduced from postulated parts, or more from postulated wholes, or more from a combination of the two? These are matters which should properly await the outcome of trial; it is conceivable that numerous distinct sets of postulates may prove more or less successful.

The history of scientific practice so far shows that, in the main, the credentials of scientific postulates have consisted in what the postulates can do, rather than in some metaphysical quibble about where they came from. If a set of postulates is really bad it will sooner or later get its user into trouble with experimental results. On the other hand, no matter how bad it looks at first; if a set of postulates consistently yields valid deductions of laboratory results, it must be good.¹¹ In a word, a complete laissez-faire policy should obtain in regard to postulates. Let the psychological theorist begin with neurological postulates, or stimulusresponse postulates, or structural postulates, or functional postulates, or factor postulates, or organismic postulates, or Gestalt postulates, or sign-Gestalt postulates, or hormonic postulates, or mechanistic postulates, or dynamic postulates, or postulates concerned with the nature of consciousness, or the postulates of dialectical materialism, and no questions should be asked about his beginning save those of consistency and the principle of parsimony.

Third, we must be extremely careful to insure the rigor of our deductions. Perhaps the most common fallacy in current would-be theories is the *non sequitur*—the supposed conclusion simply does not follow from the postulates.

¹¹ Consider the Riemannian geometry, which insists that the sum of the angles of a triangle is greater than two right angles (19, 58). This is repugnant to common sense, yet Einstein used the Riemannian geometry as the basis for making the greatest single advance in scientific theory since the time of Newton.

In particular we must be on our guard against what might be called the 'anthropomorphic fallacy.' By this is meant a deduction the critical point of which turns out to be an implicit statement which, if made explicit, would be something like. "If I were a rat and were in that situation I would do so and so." Such elements in a deduction make it a travesty because the very problem at issue is whether a system is able to deduce from its postulates alone what a normal man (or rat) would do under particular conditions. It is this fallacy which justifies the inveterate aversion of scientists for anthropomorphism. It is true that as a practical guide to the expectation of what a rat, or an ape, or a child, or another man will actually do in an as yet untried situation such an approach is, of course, of value and should be used. But predictions arrived at in such a way are of no value as scientific theory because a truly scientific theory seeks to deduce what anthropomorphism reaches by intuition or by naïve assumption. Prophecies as to the outcome of untried experiments based merely on such anthropomorphic intuitions should be credited to the intuitional genius of the prophet rather than to the theoretical system to which the prophet may adhere. Predictions, however successful, can have no evidential value as to the credibility of the prophet's system until he is willing and able to exhibit the logic by which his predictions flow from the postulates of that system, and until this logic is really rigorous, until it consists of something more than the feeble non-sequiturs too often presented in our literature as scientific explanations.

SUMMARY AND CONCLUSIONS

Scientific theory in its best sense consists of the strict logical deduction from definite postulates of what should be observed under specified conditions.¹² If the deductions are lacking or are logically invalid, there is no theory; if the deductions involve conditions of observation which are impossible of attainment, the theory is metaphysical rather than scientific; and if the deduced phenomenon is not observed

¹² Truth, for the purposes of the present paper, is to be understood as a theoretical deduction which has been verified by observation.

when the conditions are fulfilled, the theory is false. Classifications of the phenomena of a science may have distinct expository and pedagogical convenience, but convenience cannot be said to be true or false. Points of view in science may possess the virtue of fertility by suggesting new directions of investigations, but neither can fertility be said to be true or false. On the other hand, truly scientific theory, from its very nature, must permit the observational determination of its truth or falsity.

It is believed that upon the above conceptions of scientific theory may be based a robust hope of bringing order out of our present theoretical chaos. It is conceivable, of course, that more than one scientific system may be able to deduce the major phenomena of learning. However, the history of scientific theory has shown that successful duplicate explanations of the same natural phenomena have usually turned out to be at bottom the same. Accordingly, we may expect that when we have put our scientific house in order there will be little more disagreement in the field of theory than in the field of experiment, and presumably such disagreements as appear will prove to be but temporary.

Assuming both the possibility and the desirability of such an outcome, the question arises as to how it can most promptly be achieved. First, it is believed that the thing most urgently needed at the present moment is a clear statement of postulates with accompanying definitions of terms. Second. these postulates should be followed by the step-by-step deduction of the theorems making up the body of the system. No doubt the meticulous presentation of the logic behind the theorems of a system may at first strike certain readers as pedantic. Moreover, it is an unfortunate fact that for persons untrained in a particular system, the more rigorous the logic the more difficult it becomes to comprehend. It is encouraging, however, to note that difficulty of comprehension by the tyro has not prevented the development of mathematical theory in the older sciences, and with them rigor of deduction has not usually been regarded as pedantry. A number of indications point to a considerable development

of this kind of theoretical work in psychology within the immediate future.

As this development proceeds, we may anticipate that those systems or points of view which are unable to satisfy the postulational requirements of truly scientific theory will come to be known for what they are, and will lose adherents. The proponents of other points of view may be expected gradually to clarify their basic postulates and from these to evolve systems of rigorously proved theorems. Of this latter group of systems, presumably, it will be found impossible to apply the experimental check to the theorems of some because the systems in question either do not specify clearly the conditions under which phenomena should occur or else they are not clear as to exactly what phenomena are to be expected. Some systems, on the other hand, will doubtless succeed in making genuine contact with experimental facts. Of these, some will probably present such a high proportion of experimental nonconfirmations that the confirmations actually observed may be attributable to mere chance.

Finally, let us hope, there will survive a limited number of systems which show a degree of successes appreciably in excess of what chance would produce. Occasionally, in such cases, a failure of a theorem to agree with experimental observation may be accounted for plausibly on the basis of a known and recognized factor operating in such a way as to over-ride the action represented by the theorem. Unless this can be done, however, the postulates of the system must be revised until they yield theorems agreeing with both the new and the old facts, after which there will be made new deductions which will be checked against new experiments, and so on in recurring cycles. Thus theoretical truth is not absolute, but relative.

It seems likely that as the process of theoretical development goes on the surviving systems will show two fairly distinct types of relationship. First, there will be systems which attempt explanations on different levels such as the perceptual level, the stimulus-response level, the neuro-anatomical level, and the neuro-physiological level. It is conceivable that each might develop a perfect system on its own level. In that case each lower level should be able to deduce the relevant basic postulates of the system above it in the hierarchy of systems. Here, of course, would be supplementation rather than conflict.

Second, there may be some systems which attempt explanation at the same level. However diverse such systems may appear at the beginning, they may be expected gradually to display an essential identity as they go through successive revisions, the differences at length consisting in nothing but the terms employed. Those systems which concern different but related aspects of learning, by the process of expansion, will finally come to overlap. This overlapping will convert them into approximately the same status as the groups just mentioned, and a gradually approached outcome of substantial agreement may similarly be anticipated. Thus systems may expand by a process of integration.

Finally, sound scientific theory has usually led not only to prediction but to control; abstract principles in the long run have led to concrete application. With powerful deductive instruments at our disposal we should be able to predict the outcome of learning not only under untried laboratory conditions, but under as yet untried conditions of practical education. We should be able not only to predict what rats will do in a maze under as yet untried circumstances, but what a man will do under the complex conditions of everyday life. In short, the attainment of a genuinely scientific theory of mammalian behavior offers the promise of development in the understanding and control of human conduct in its immensely varied aspects which will be comparable to the control already achieved over inanimate nature, and of which the modern world is in such dire need.

REFERENCES

- I. ADAMS, G., Psychology: science or superstition?, New York; Covici Friede, 1931.
- 2. BACON, F., Advancement of learning and novum organum, Revised edition, Colonial Press, 1900.
- 3. CROZIER, W. J., Chemoreception, from A handbook of general experimental psychology, edited by Carl Murchison, Worcester: Clark University Press, 1934.

515

- 4. EBBINGHAUS, H., Memory (trans. by H. A. Ruger and C. E. Bussenius), N. Y.: Teachers College, Columbia University, 1913.
- 5. EVANS, C. L., Recent advances in physiology, Phila.: P. Blakiston's Son and Co., 1926.
- 6. GULLIKSEN, H. A., A rational equation of the learning curve based on Thorndike's law, J. General Psychol., 1934, 11, 395-434.
- 7. HAAS, A., Introduction to theoretical physics, Vol. II (trans. from the 3rd and 4th editions by T. Verschoyle), London: Constable and Co., 1925.
- 8. HECHT, S., Vision: II. The nature of the photoreceptor process, from A handbook of general experimental psychology, edited by Carl Murchison, Worcester: Clark University Press, 1934.
- 9. HULL, C. L., The influence of tobacco smoking on mental and motor efficiency, Psychol. Monog., 1924, 33, No. 150.
- The influence of caffeine and other factors on the phenomena of rote learning, J. General Psychol., 1935, 13, 249-274.
- KREUGER, W. C. F., Effect of overlearning on retention, J. Exper. Psychol., 1929, 12, 71-78.
- 12. LEPLEY, W. M., A theory of serial learning and forgetting based upon conditioned reflex principles, PSYCHOL. REV., 1932, 39, 279-288.
- ----, Serial reactions considered as conditioned reactions, Psychol. Monog., 1934. No. 205.
- 14. MURCHISON, C., (Editor), Psychologies of 1925, Worcester: Clark University Press, 1925.
- 15. ----, Psychologies of 1930, Worcester: Clark University Press, 1930.
- Newton, Isaac, Mathematical principles (Trans. by F. Cajori), Univ. of California Press, 1934.
- 17. PAVLOV, I. P., Conditioned reflexes (trans. by G. V. Anrep), Oxford University Press, 1927.
- 18. PILLSBURY, W. B., The fundamentals of psychology, N. Y.: Macmillan, 1927.
- 19. POINCARÉ, H., The foundations of science (trans. by G. B. Halsted), The Science Press, 1929.
- 20. REICHE, F., The quantum theory (trans. by H. S. Hatfield and H. L. Brose), N. Y.: E. P. Dutton and Co., 1930.
- 21. ROBINSON, E. S., AND BROWN, M. A., Effect of serial position on memorization, Amer. J. Psychol., 1926, 37, 538-552.
- 22. SHERRINGTON, C. S., The integrative action of the nervous system, New Haven: Yale University Press, 1906.
- 23. SIMLEY, O. H., The relation of subliminal to supraliminal learning, Arch. Psychol., 1933, No. 146.
- 24. WARD, L. B., Retention over short intervals of time, Thesis presented for the degree of Doctor of Philosophy, Yale University, 1934 (on file in the Yale University Library).
- [MS. received April 12, 1935]